72P

INVESTIGATION OF SUFFICIENCY CONDITIONS AND THE HAMILTON JACOBI APPROACH TO OPTIMAL CONTROL PROBLEMS

MARTIN-CR-64-1 (Issue 7) Copy No. 2

by

HENRY G. HERMES

APPLIED MATHEMATICS SECTION
ELECTRONICS AND MATHEMATICS DEPARTMENT
MARTIN COMPANY, DENVER, COLORADO

FINAL REPORT

JULY 1964

CONTRACT NAS8-11020

prepared for

GEORGE C. MARSHALL SPACE FLIGHT CENTER
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
HUNTSVILLE, ALABAMA

INVESTIGATION OF SUFFICIENCY CONDITIONS AND THE HAMILTON JACOBI APPROACH TO OPTIMAL CONTROL PROBLEMS

By

H. Hermes

Applied Mathematics Section
Electronics Laboratory
Martin Co., Denver, Colorado

July 1964

Prepared with the support of NASA contract NAS 8-11020, George C. Marshall Space Flight Center, Huntsville, Alabama.

TABLE OF CONTENTS

INTRODUCTION AND SUMMARY

This is the final report on contract NAS 8-11020 entitled "Optimum Trajectory Study".

In this section we will try to give a verbal account of the problems considered, the reasons for considering them, and the main results obtained. The remaining sections, while having independent introductions, will contain the mathematical analysis.

The major objective of this study was to examine the use of Hamilton Jacobi partial differential equations in determining fields of optimum trajectories and to study sufficiency conditions. Since a great number of optimal control problems can, with a slight reformulation, be posed as time optimal problems, our attention is focused throughout on problems of this type.

If given initial data, say time $t = t_o$, state $x = x_o$ for a time optimal problem, the reachable set (in Euclidean (n+1) dimensional time—state space) is defined to be the set of all points (t, x) with time $t \geq t_o$ and state x such that it can be attained in time t by a trajectory of the dynamical system with an admissible control. Under very mild conditions on the dynamical system equations and the control set, it is known that a time optimal point to point transfer will lead to a trajectory which lies on the boundary of the reachable set. Conversely, trajectories which lie on the boundary of the reachable set are excellent candidates for being time optimal for some point to point transfer, and thus conditions which single them out are of interest. Now a point is on the boundary of the reachable set if in every neighborhood of it there are points not in the reachable set; i.e., points not attainable by trajectories of the dynamical system. This leads one naturally to notions of controllability.

Following the definition of Kalman, a linear system is said to be completely controllable at time t_o if every state can be attained (with $\mathcal{L}_2$ control) in finite time by a trajectory of the system having arbitrary initial data (t_o, x_o). Thus one can examine whether the terminal data has been chosen so that the mission is possible. It is of further interest to define local controllability, i.e., a system is locally controllable along a solution trajectory $\wp(t)$ if for some $t_1 > t_o$ all points in some state space neighborhood of $\wp(t_1)$ are attainable in time t_1 by trajectories with admissible controls. Obviously trajectories along which a system is locally controllable cannot remain on the boundary of the reachable set, and hence this becomes a test for optimality. It might also be remarked that while for linear systems one could expect global controllability results, for nonlinear systems it is natural to expect only local results.

In Section I, the Kalman criterion for complete controllability for a linear system is derived in a simple manner (corollary I.1) and an extension is obtained for a special form of nonlinear system (Theorem I.2).

In Section II, the nonlinear system $\dot{x}(t) = g(t, x(t))+H(t,x(t))u(t)$, x an n vector, H an nxr matrix, u and r vector valued control with $1 \leq r < n$, is studied. If $B(t, x)$ is an $(n-r)xn$ matrix, of maximal rank, such that $B(t,x)H(t,x) \equiv 0$, the local controllability of the above system is shown to be closely related to the integrability of the pfaffian system $B(t,x)dx - B(t,x)\, g(t, x)dt = 0$. In particular, the above nonlinear system is defined to be completely controllable if the associated pfaffian system is not integrable. Theorem II.1 then shows that in the special case of a linear system, this definition yields a criterion for complete controllability equivalent to that of Kalman. This new criterion is useful since it does not depend on the knowledge of a fundamental solution matrix for a time varying linear system. Its use is demonstrated by obtaining the result that an n dimensional system, formed from a single n^{th} order linear time varying differential equation of the form $x^{(n)}(t) + a_1(t)\ x^{(n-1)}(t) + \ldots + a_n(t)\ x(t) = u(t)$, is completely controllable. (Here u is a scalar valued control). This result was previously known if the functions $a_i(t)$ were constant.

The remainder of section II deals with local controllability in a neighborhood of singular arcs. It is shown that local tests, which depend on examining the controllability of the variational equation along a singular arc will always be non-conclusive. Along an optimal singular arc the system is truly not locally controllable, however it is shown by example (example II.2) that singular arcs can exist along which the system is locally controllable. These can be thought of as inflection points in function space, of the functional (time) which is to be extremized. They are analogous to inflection points which arise when extremizing a real valued function F on a manifold in Euclidean space; i.e., non-extremal points at which the map F induces on the tangent space of the manifold into the tangent space of the reals, vanishes.

These arcs are singular also in the sense of the classical calculus of variations, hence the Hilbert differentiability condition fails to hold along them, and classical sufficiency conditions fail.

In section III, the study of feedback control via the Pontriagin maximum principle and Hamilton Jacobi theory is begun. Often the feedback control which the maximum principle prescribes, is discontinuous in the state variables, which in turn leads to a Hamilton Jacobi equation with discontinuous coefficients. This is impractical both from a theoretical and computational viewpoint. The first part of section III deals mainly with the reason for this discontinuity, and yields conditions such that the maximum principle would prescribe a continuous or even C^1 (once continuously differentiable) control. Theorems III.4 and III.5 then show that whenever a control problem merely satisfies the conditions of Fillipov for the existence of an optimal control, there exists an approximate problem (the precise definition of this precedes theorem III.4) for which the maximum principle gives a C^1 control, and such that for any given $\epsilon > 0$, an optimal trajectory of the original problem will be in an ϵ neighborhood of that for the approximate problem.

The remainder of section III deals with the Hamilton Jacobi theory for these smooth approximate problems, and for the special case of the control appearing linearly, an easy construction for the approximating problem is shown, while an example (example III.1) is worked out in detail to demonstrate the results.

Two sets of references are given, the first for sections I and II, the second for section III.

CONTROLLABILITY AND THE SINGULAR PROBLEM

INTRODUCTION TO SECTIONS I AND II

The concept of complete controllability of linear systems was introduced by R. E. Kalman [1]. It is part of the purpose of this paper to extend the concept to nonlinear systems, with control appearing linearly. All systems considered are of this form.

Geometrically, a linear system is completely controllable at time t_o if any state can be attained in finite time by a trajectory of the system having arbitrary initial data x_o at time t_o. The motivation for the extension of this concept to nonlinear systems came largely from results obtained in [2] and from the geometric interpretation of non-integrability of pfaffians given in [3] and [4]. In particular, Caratheodory gives an argument to show that if, for a single pfaffian equation, there are points in every neighborhood of a given point which are not "reachable" from the given point by curves satisfying the equation, the equation is integrable. This result was generalized to systems of pfaffians in [4]. There is a difficulty in applying these ideas to pfaffian systems which are quite naturally associated with control systems having control appearing linearly. (See § II.) The reason for this is that usually the independent variable t appears explicitly in the pfaffian system, hence its integral curves, which can be related back to solutions of the control system, and are used to connect neighboring points to a given point, must have t parametrized as $t(\sigma)$, a monotone function of σ. This is not the case in the proofs in [3] and [4], and with this restriction, in general the results of these papers are no longer valid.

The relation between singular problems and controllability arises quite naturally from the pfaffian approach and can be anticipated from results obtained by LaSalle in [5]. In § II we define the concept of a totally singular arc, i.e., an arc satisfying the differential constraining equations, for which there exists an adjoint vector such that the maximum principle yields no information as to the optimality of any of the components of the control along this arc. In particular, if the system were linear and admitted no totally singular arc, the system would be proper in the sense of LaSalle [5] and completely controllable in the sense of Kalman [6]. Even if the controls are merely restricted to be $\mathcal{L}_2$ (Lebesgue square integrable) functions, it is shown that totally singular arcs can exist and comprise some or all of the boundary of the attainable set, thereby being optimal trajectories for certain time optimal control problems. These are also precisely the arcs along which the system need not be locally controllable, i.e., if we assume initial data x_0 given at time t_0, there may exist points in every state space neighborhood of a point $\varphi^v(t_1)$ of a totally singular arc φ^v, which are not attainable in time $t_1 > t_0$ by trajectories of the system with $\mathcal{L}_2$ controls. Here φ^v denotes the solution of the system with control v. Precisely, if for every $t_1 > t_0$ there exist points in every state space neighborhood of $\varphi^v(t_1)$, which are not attainable with $\mathcal{L}_2$ control in time t_1, the arc φ^v is totally singular. However it is shown by example that there do exist totally singular arcs about which the system is locally controllable.

§ I. COMPLETE CONTROLLABILITY FOR LINEAR AND MILDLY NONLINEAR SYSTEMS

Throughout this section H will denote an nxr matrix valued function of t, which is in $\mathcal{L}_2 \left[t_o, t_1\right]$ for any given finite $t_1 > t_o$. Controls will be $\mathcal{L}_2$, vector valued functions, We begin with the following basic Lemma.

<u>Lemma I.1</u> A necessary and sufficient condition that there exist an rxn matrix valued function V(t) in $\mathcal{L}_2 \left[t_o, t_1\right]$, such that for some $t_1 > t_o$, $\int_{t_o}^{t_1} H(\tau)V(\tau)d\tau$ is non-singular, is that for some $t_1 > t_o$ $\int_{t_o}^{t_1} H(\tau)H^T(\tau)d\tau$ is non-singular.

<u>Proof</u> Sufficiency is immediate by choosing $V(\tau) = H^T(\tau)$. To show necessity assume there exist V, $t_1 > t_o$, such that $\int_{t_o}^{t_1} H(\tau)V(\tau)d\tau$ is non-singular, but $\int_{t_o}^{\bar{t}} H(\tau)H^T(\tau)d\tau$ is singular for all $\bar{t} > t_o$, in particular $\bar{t} = t_1$. This implies there exists a constant vector $c \neq 0$ such that $c\left(\int_{t_o}^{t_1} H(\tau)H^T(\tau)d\tau\right) c^T = 0$, and since $H(\tau)H^T(\tau)$ is positive semi-definite, we obtain $cH(t) = 0$ almost everywhere in $\left[t_o, t_1\right]$. Thus

$\int_{t_o}^{t_1} H(\mathcal{T})V(\mathcal{T})d\mathcal{T} = 0$ which contradicts the non-singularity of

$\int_{t_o}^{t_1} H(\mathcal{T})V(\mathcal{T})d\mathcal{T}$. ∎

We next consider the system

$$(1\text{-}1) \qquad \dot{x}(t) = H(t)u(t) \ , \quad x(t_o) = x_o, \quad u \in \mathcal{L}_2 \left[t_o, \ t_1\right].$$

Define

$$M(t_o, \ t_1) \equiv \int_{t_o}^{t_1} H(\mathcal{T})H^T(\mathcal{T})d\mathcal{T} \ .$$

<u>Theorem I.1</u> A necessary and sufficient condition for the system (1-1) to be completely controllable at t_o is that there exists $t_1 > t_o$ such that $M(t_o, \ t_1)$ is non-singular.

<u>Proof:</u> (Sufficiency) Let $\overline{x}$ be any given point in E^n, Euclidean n space. We will show $\overline{x}$ is attainable from x_o at time t_1. Indeed pick $u(t) = H^T(t)\xi$, $\xi \in E^n$. We desire $\overline{x} = x(t_1) = x(t_o) + \left(\int_{t_o}^{t_1} H(\mathcal{T})H^T(\mathcal{T})d\mathcal{T}\right)\xi$ or $\xi = M^{-1}(t_o,t_1) \ (\overline{x} - x(t_o))$.

(Necessity). Assume $M(t_o, \ t_1)$ is singular for all $t_1 > t_o$. This implies (see proof of lemma I.1) that there exists a constant vector $c \neq 0$ such that $c \ H(t) \equiv 0$ p.p. Since x_o is arbitrary, let it be such that $c \circ x_o = 0$. We will show the point c is not attainable from x_o. Indeed suppose for some u and t_1, $c = x_o + \int_{t_o}^{t_1} H(\mathcal{T})u(\mathcal{T})d\mathcal{T}$. Then

$c \cdot c = \|c\|^2 = c \cdot x_o + c\int_{t_o}^{t_1} H(\tau)u(\tau)d\tau = 0$, a contradiction to the fact that $c \neq 0$. ∎

<u>Corollary I.1</u> (Kalman) The linear system

(1-2) $\quad \dot{x}(t) = A(t)x(t) + H(t)u(t) , \qquad x(t_o) = x_o$

is completely controllable at t_o if and only if

$\int_{t_o}^{t_1} \Phi(t_o,\tau)H(\tau)H^T(\tau)\Phi^T(t_o,\tau)d\tau$ is non-singular for some $t_1 > t_o$.

Here $\Phi(t,\tau)$ denotes a fundamental solution of the homogeneous system $\dot{x}(t) = A(t)\,x(t)$.

<u>Proof</u>: Make the transformation $y(t) = \Phi^{-1}(t, t_o)\,x(t)$. Then x satisfies (1-2) if an only if y satisfies

(1-3) $\quad \dot{y}(t) = \Phi(t_o,t)\,H(t)u(t), \quad y(t_o) = x_o.$

(Note $\Phi(t_o, t) = \Phi^{-1}(t, t_o)$.) From the transformation, it follows that the system (1-2) is completely controllable if and only if the system (1-3) is completely controllable, i.e., from theorem I.1 that there exists a $t_1 > t_o$ such that

$\int_{t_o}^{t_1} \Phi(t_o,\tau)\,H(\tau)\,H^T(\tau)\,\Phi^T(t_o,\tau)d\tau$ is non-singular. ∎

Some special results for nonlinear systems

We next consider the nonlinear system

(1-4) $\dot{x}(t) = g(t, x(t)) + H(t)u(t), \qquad x(t_o) = x_o$

with the assumptions: i) $\left| g^j(t,x) \right| \leq M$, $j = 1, 2, \ldots, n$.
ii) $\left| g^j(t,x) - g^j(t, \bar{x}) \right| \leq m \, \| x - \bar{x} \|$, $j = 1, 2, \ldots, n$. iii) g is continuous as a function of t for each x.

Again let $M(t_o,t_1) = \int_{t_o}^{t_1} H(\tau)H^T(\tau)d\tau$.

Theorem I.2 A sufficient condition that the set of points attainable by trajectories of the system (1-4) with $\mathcal{L}_2$ control be all of E^n is that $M(t_o, t_1)$ be non-singular for some $t_1 > t_o$.

Remark Rather than state the theorem in this manner, one might consider merely saying that the system (1-4) is completely controllable at t_o. However, this notion has not been defined for nonlinear systems, and it does not seem reasonable to this author to define it in such a global fashion for these systems.

Proof For arbitrary u, (1-4) has a solution designated φ^u which satisfies

$$(1\text{-}5) \quad \varphi^u(t) \equiv x_o + \int_{t_o}^{t} g(\tau, \varphi^u(\tau))d\tau + \int_{t_o}^{t} H(\tau)u(\tau)d\tau \ .$$

Let $\bar{x}$ be any given point in E^n. We desire a control such that for some point finite $t_1 > t_o$, $\varphi^u(t_1) = \bar{x}$. It suffices to consider controls which come from a finite dimensional subspace of $\mathcal{L}_2$, in particular the controls considered will be of the form $u(t) = H^T(t)\xi$ where $\xi \in E^n$. Hence the notation φ^{ξ} rather than φ^u will be used.

Define a mapping $\mathcal{F}: E^n \longrightarrow E^n$ as follows:

Let $\alpha(\xi) \equiv \int_{t_o}^{t_1} g(\tau, \varphi^{\xi}(\tau))d\tau$, and define

$\mathcal{F}(\xi) \equiv M^{-1}(t_o, t_1)\left[\bar{x} - \alpha(\xi) - x_o\right]$. From (1-5) it follows that a fixed point of $\mathcal{F}$ will yield a value ξ such that $\varphi^{\xi}(t_1) = \bar{x}$.

It is well known that with the conditions imposed on g [7, th. 7.4 - Chapter I], φ^{ξ} is a continuous function of ξ in the topology $C[t_o, t_1]$, i.e., the topology induced by the supremum norm. Thus $\alpha(\xi)$ is a continuous function of ξ, and $\mathcal{F}$ is a continuous function of ξ.

We next show that there exists a K such that $\|\xi\| \leq K \longrightarrow \|\mathcal{F}(\xi)\| \leq K$. Letting $\|\xi\| = \sum_{i=1}^{n} |\xi_i|$, and $\|M^{-1}\|$ be any matrix norm, since $|g^j| \leq M$, for any ξ, $\|\alpha(\xi)\| \leq n(t_1 - t_o)M$:

Letting $K \equiv \|M^{-1}(t_o, t_1)\| \left[\|\bar{x}\| + nM(t_1 - t_o) + \|x_o\|\right]$, it follows that for any ξ, $\|\mathcal{F}(\xi)\| \leq K$, hence in particular $\mathcal{F}$ maps the ball $\left\{\xi \in E^n: \|\xi\| \leq K\right\}$ continuously into itself. Thus $\mathcal{F}$ has a fixed point. ∎

Remark The result obtained in this theorem is not surprising in view of theorem (I.1) and the boundedness condition on the vector g. Also the condition $M(t_o, t_1)$ non-singular for some $t_1 > t_o$ is much stronger than it need be. For example, if we consider a linear system of the form (1-2) and $H(t)$ is a column vector with one component zero, then $M(t_o,t_1)$ is singular for all $t_1 \geq t_o$, yet the system can certainly be completely controllable.

§ II. NONLINEAR SYSTEMS WITH LINEAR CONTROL; THE SINGULAR PROBLEM

In this section, we consider extending the notion of complete controllability to systems of the form

$$\dot{x}(t) = g(t, x(t)) + H(t, x(t))u(t) \tag{2-1}$$

where g is an n-vector, H an nxr matrix, while u is an $\mathcal{L}_2$ control vector. It is assumed that g and H are C^1 in all arguments. Throughout, the stipulation $1 \leq r < n$ is required to hold.

Let $B(t, x)$ be a C^1, (n-r)xn matrix with rank (n-rank H) at each point (t, x) in some domain $\mathcal{D}$, of interest, such that

$$B(t,x)\, H(t, x) \equiv 0 , \qquad (t, x) \in \mathcal{D}. \tag{2-2}$$

Since $r < n$, we know that rank $B \geq 1$ for all (t, x).

With the system (2-1), associate the pfaffian system

$$B(t, x)dx - B(t, x)\, g(t, x)dt = 0. \tag{2-3}$$

Let b be an arbitrary linear combination of the rows b^{ν} of B, taken with C^1 scalar valued coefficients $\alpha_{\nu}(t, x)$, i.e.,

$b(t, x) = \sum_{\nu} \alpha_{\nu}(t, x)\, b^{\nu}(t, x)$. Throughout, b will be used to denote such a linear combination which is not identically zero.

Definition II.1 The pfaffian system (2-3) is integrable at the point $(\bar{t}, \bar{x})$ if there exists a C^1 scalar valued function $\psi(t, x)$ and an $\epsilon > 0$ such that for some b,

$$\psi_x(t, x) = b(t, x), \quad \psi_t(t, x) = -b(t, x) \cdot g(t, x)$$

for $\bar{t} \leq t < \bar{t} + \epsilon$, $|x - \bar{x}| < \epsilon$.

Essentially this states that for some b,

(2-4) $b(t, x)dx - b(t, x) \cdot g(t, x)dt$

is an exact differential in a "neighborhood" of $(\bar{t}, \bar{x})$. It should be noted that any integrating factor can be included in the coefficients of the linear combination of the rows b^{ν}.

The notion of integrability of a pfaffian system is, of course, related to the property of completencess of an associated system of partial differential equations. To show the relation, let $C(x)$, $x \in E^n$, be a smooth (n-r)xn matrix, and $K(x)$ a smooth nxr matrix, both of maximum rank, such that $C(x)K(x) \equiv 0$. With the pfaffian system

(2-4) $C(x)dx = 0$

we associate the system of partial differential equations $K^T(x) \dfrac{\partial f(x)}{\partial x} = 0$.

Each row k^i of K^T can be considered as defining a vector field X^i which locally generates a one parameter semi group of diffeomorphisms, $\{T_i(t)\}$, see for example [8, p. 10]. In turn, such a semi group determines a vector field. If for each i, j = 1, 2, ..., r and for all arbitrarily small fixed τ, the vector field determined by $\{T_j(\tau)\ T_i(t)\ T_j(-\tau)\}$ is linearly dependent on the fields X^i, the system of partial differential equations is said to be complete. If it is not complete, the number m of linearly independent fields formed in this manner is called the index of both the pfaffian system and the associated partial differential equation system [4].

From the results in [4], it easily follows that <u>the pfaffian system (2-4) is integrable (definition II.1) if and only if the index m is such that $m+r < n$</u>. If the index m is such that $m+r = n$, Chow [4] shows that there is a neighborhood of a point $x_o \in E^n$ such that all points in this neighborhood are attainable by curves satisfying (2-5). From the viewpoint of local controllability for a control system, we can interpret this as follows. <u>If the pfaffian system associated with the control system</u>

$$(2\text{-}5)\qquad \dot{x}(t) = K(x(t))u(t)\ , \qquad\qquad x(t_o) = x_o$$

<u>has index m, where K is a continuous nxr matrix function of $x \in E^n$ with constant rank r, and $m+r = n$, then every point in some neighborhood of x_o is attainable by trajectories of (2-5) with measurable controls</u>. Indeed, since all points in some neighborhood of x_o are attainable by

absolutely continuous curves satisfying $C(x,(t))\,\dot{x}(t) = 0$ almost everywhere, we must only show that such a curve also satisfies (2-5) for some control u. But $C(x(t))\,\dot{x}(t) = 0 \Longrightarrow \dot{x}(t)$ is a linear combination of the columns of $K(x(t))$, since $CK \equiv 0$. Thus there exists $u(t)$ such that $\dot{x}(t) = K(x(t))u(t)$ for almost all t. Since K has rank r, it has a continuous left inverse on its range, from which it follows that u is measurable.

Before stating an explicit criterion for complete controllability of a system of the form (2-1) one may ask: <u>What should one expect the definition to yield</u>? This can presently be answered as follows. Since the definition should extend that given for a linear system of the form (1-2) which is a special case of (2-1), one expects:

a) If $g(t, x) \equiv A(t)x$, $H(t, x) \equiv H(t)$, then the criterion which defines complete controllability at t_o for (II.1) should be equivalent with the condition $\int_{t_o}^{t_1} \Phi(t_o, t)H(t)H^T(t)\,\Phi^T(t_o, t)dt$ non-singular for some $t_1 > t_o$, as given in corollary I.1.

b) There should be a geometric interpretation of the condition, e.g., what points are attainable from the initial point in finite time? In the linear system there were global attainability results, i.e., any point could be attained from the initial point via a trajectory of the system. In the nonlinear problem, one would expect at most local results of this nature.

The approach will be to state a criterion for complete controllability of (2-1) which we will show satisfies a). We then use this criterion to try to establish a geometric interpretation as mentioned in b). Of course, how the definition of complete controllability should be extended is somewhat a matter of personal opinion.

Definition II.2 The system (2-1) is completely controllable at $(\bar{t}, \bar{x}) \in \mathscr{D}$ if the associated pfaffian system (2-2) is not integrable at $(\bar{t}, \bar{x})$.

It will next be shown that this criterion is equivalent to the condition given in corollary I.1 for the special case of the linear system (1-2). In this case it suffices to take $B = B(t)$ in forming the pfaffian system equivalent to (2-3). Also, in taking the linear combination of the rows of B to form the single pfaffian as in (2-4), we can consider the scalar functions α_ν as function of only t. Indeed we must only show that if the pfaffian form

(2-6) $\quad b(t)dx - b(t)\,A(t)x\,dt$

has an integrating factor, then this integrating factor, denoted by μ, can be taken as a function of only t. To obtain this, suppose $\bar{\mu}(t, x)$ is such that $\bar{\mu}(t, x)\,b(t)dx - \bar{\mu}(t, x)b(t)A(t)x\,dt$ is an exact differential. Then $\bar{\mu}_{x_j}\,b^i - \bar{u}_{x_i}\,b^j = 0$ for all i, j = 1, 2, ..., n, and $\bar{\mu}_t b + \bar{\mu}\dot{b} = -\bar{\mu}_x\,b\,A\,x - \bar{\mu}\,b\,A$. Define $\mu(t) = \bar{\mu}(t, 0)$, noting that for the linear system $\mathcal{D} = (t_o, \infty) \times E^n$ which implies $(t, 0) \in \mathcal{D}$ for $t > t_o$. It follows that $\mu(t)$ is also an integrating factor.

Since it is sufficient to consider both μ and the α_ν as functions of only t, there is no loss of generality in considering that if the pfaffian system

$$(2\text{-}7) \qquad B(t)dx - B(t)\, A(t)x\, dt = 0$$

associated with (1-2) is integrable, then (2-6) is an exact differential.

Since x appears linearly, definition II.1 simplifies for such systems, and is: <u>The pfaffian system (2-7) is integrable at the point $\bar{t}$ if there exists a C^1 scalar valued function $\Psi(t, x)$ and an $\epsilon > 0$ such that for some b,</u>

$$\Psi_x(t, x) = b(t), \quad \Psi_t(t, x) = -b(t)\, A(t)\, x$$

<u>for $\bar{t} \leq t < \bar{t} + \epsilon$</u>. (Note: Under the assumptions on B and H, Ψ_{xt} and Ψ_{tx} exist and are equal).

Define:

$$W(t_o, t_1) = \int_{t_o}^{t_1} \Phi(t_o, t)H(t)H^T(t)\, \Phi^T(t_o, t)dt \,.$$

Then corollary I.1 states that the system (1-2) is completely controllable at t_o if and only if there exists a $t_1 > t_o$ such that $W(t_o, t_1)$ is non-singular.

<u>Remark 1</u>. If A and H are constant matrices, Kalman [1] shows that this condition is equivalent to the condition: rank $[A, AH, \ldots A^{n-1}H] = n$.

Remark 2. While the above condition given for the constant coefficient case can be directly checked, $W(t_o, t_1)$ depends on knowledge of a fundamental solution $\Phi(t, t_o)$ which is not always easily obtainable.

Remark 3. It is easily seen that $W(t_o, t_1)$ is a positive semi-definite matrix. Thus if $W(t_o, t_1)$ is non-singular, $W(t_o, t)$ is non-singular for all $t \geq t_1$.

The main purpose of this section will be to show that the condition II.2 for complete controllability of (1-2) is equivalent to $W(t_o, t_1)$ being non-singular for some $t_1 > t_o$. This condition has the advantage of not depending on knowledge of a fundamental solution.

Before stating the main theorem, a simple computation yields, for $t_o < t_1 < t_2$,

$$W(t_o, t_2) = W(t_o, t_1) + \Phi(t_o, t_1)\, W(t_1, t_2)\, \Phi^T(t_o, t_1).$$

Thus if $W(t_1, t_2)$ is non-singular (positive definitive) it follows that $W(t_o, t_2)$ is also non-singular (positive definite). The reverse implication need not be true.

Theorem II.1 A necessary and sufficient condition that $W(t_1, t_2)$ be non-singular for all $t_2 > t_1$ is that the pfaffian (2-7) be not integrable at t_1.

For ease in both using and proving this theorem, we list the implications and their contrapositives.

I.A <u>Necessary condition</u>: $W(t_1, t_2)$ non-singular for all $t_2 > t_1$

$\Longrightarrow$ pfaffian (2-7) is not integrable at t_1.

I.B <u>Necessary; contrapositive</u>: Pfaffian (2-7) integrable at $t_1 \Longrightarrow$

$W(t_1, t_2)$ is singular for some $t_2 > t_1$.

I.C <u>Sufficient condition</u>: Pfaffian (2-7) not integrable at t_1

$\Longrightarrow W(t_1, t_2)$ is non-singular for all $t_2 > t_1$.

I.D <u>Sufficient; contrapositive</u>: $W(t_1, t_2)$ singular for some $t_2 > t_1$

$\Longrightarrow$ pfaffian (2-7) is integrable at t_1.

<u>Proof</u>: (We shall prove I.B and I.D)

Assume the pfaffian (2-7) is integrable at t_1. Then there is a vector b, which is a linear combination of the rows of B, and an $\epsilon > 0$ such that $\dot{b}(t) = -b(t)A(t)$, for $t_1 \leq t \leq t_1 + \epsilon$. Let $\Phi(t, t_1)$; $\Phi(t_1, t_1) = I$, be the fundamental solution of $\dot{x} = A(t)\, x$. Then the vector b admits the representation $b(t) = c\, \Phi^{-1}(t, t_1) = c\, \Phi(t_1, t)$ for some constant vector c. Let h(t) be any column of H(t). Then $0 = b(t)h(t) = c\, \Phi(t_1, t)\, h(t)$. Since h was an arbitrary column of H, and W is positive semi-definite, we have $c\, W(t_1, t)\, c^T = 0$ for $t_1 \leq t \leq t_1 + \epsilon$ showing that there exists a $t_2 > t_1$ such that $W(t_1, t_2)$ is singular.

Assume, next, that $W(t_1, t_2)$ is singular for some $t_2 > t_1$. From remark 3, it follows that $W(t_1, t)$ is singular for all $t_1 \leq t \leq t_2$. This implies there exists a vector $c(t_2)$ such that $c(t_2)W(t_1, t_2)c^T(t_2) = 0$.

Since the integrand of the integral defining $W(t_1, t_2)$ is continuous,

$$c(t_2)\, \Phi(t_1, t)H(t)H^T(t)\, \Phi^T(t_1, t)c^T(t_2) \equiv 0 \quad \text{for } t_1 \leq t \leq t_2.$$

It follows that $0 \equiv c(t_2)\, \Phi(t_1, t)H(t) \equiv c(t_2)\, \Phi^{-1}(t, t_1)H(t)$, thus b defined by $b(t) \equiv c(t_2)\, \Phi^{-1}(t, t_1)$ is an admissible vector in the sense that $b(t)\, H(t) \equiv 0$, i.e., b lies in the subspace spanned by the rows of B.

Define the scalar valued function $\Psi(t,x) = c(t_2)\, \Phi^{-1}(t, t_1)x$. Then $\Psi_x(t, x) = b(t)$, $\Psi_t(t, x) = -b(t)\, A(t)x$ for $t_1 \leq t \leq t_2$ showing that the pfaffain (2-7) is integrable at t_1. $\blacksquare$

The following illustrates the advantage of a definition of complete controllability for linear systems which does not depend on knowledge of a fundamental solution.

It is known that an n dimensional system which is formed from a single n^{th} order equation having constant coefficients and the control as forcing term is completely controllable. We next show that this is also true for time varying systems of the form

$$x^{(n)}(t) + a_1(t)\, x^{(n-1)}(t) + \dots + a_n(t)\, x(t) = u(t).$$

Specifically we shall show that for any t_o, the associated pfaffian is not integrable implying $W(t_o, t_1)$ is non-singular for all $t_1 > t_o$.

We take the equivalent linear system of the form

$\dot{y}(t) = A(t)\, y(t) + h(t)\, u(t)$ where

$$A(t) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 0 \\ 0 & & & 0 & 1 \\ -a_n, & -a_{n-1}, & \cdots, & & -a_1 \end{bmatrix}, \qquad h(t) = \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

One can choose $B(t)$ as the $(n-1) \times n$ matrix

$$B(t) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & & & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} .$$

The pfaffian system equivalent to (2-7) is then

$$\text{(2-8)} \qquad \begin{aligned} dx_1 - x_2\,dt &= 0 \\ dx_2 - x_3\,dt &= 0 \\ \vdots \qquad & \\ dx_{n-1} - x_n\,dt &= 0. \end{aligned}$$

If (2-8) were to be integrable there must exist scalar valued functions $\alpha_j(t)$, not all zero, so that the single pfaffian

$$\sum_{j=1}^{n-1} \alpha_j(t)\,dx_j + 0\,dx_n - \sum_{j=1}^{n-1} \alpha_j(t)\,x_{j+1}\,dt$$

is an exact differential. But this would imply $\alpha_j(t) = 0$, $j = 1, 2, \ldots, (n-1)$, which shows (2-8) is not integrable for any t_o.

Geometric Interpretation, Local Controllability, and the Singular Problem

By associated a pfaffian system of the form (2-3) with the system (2-1), it is conspicuous that the stress is taken away from the functional form of the elements of the matrix H, and placed only on what the range of $H(t, x)$, considered as an operator on E^r, is. This obviously should be the case when controls are required to be only $\mathcal{L}_2$ functions.

In [9], Markus and Lee consider a system of the form $\dot{x} = f(x, u)$, $f \in C^1$ in $E^n \times \Omega$, where Ω a compact set contained in E^r with 0 in its interior, is the range set of the control. Assuming $f(0, 0) = 0$ and letting $A = f_x(0, 0)$, $H = f_u(0, 0)$, it is shown that if the linear system $\dot{x} = Ax + Hu$ is completely controllable, then the set of points from which the origin can be reached in finite time by trajectories of $\dot{x} = f(x, u)$, is an open connected set containing the origin. Kalman [10] pointed out that a similar result can be obtained for a system of the form $\dot{x} = f(t, x, u)$ by assuming the linear approximation is completely controllable in terms of the criterion given in corollary I.1.

The system

$$(2\text{-}9) \qquad \dot{x}(t) = f(t, x(t), u(t)) \quad , \qquad x(t_o) = x_o$$

where x is an n vector, f is a C^2 vector valued function and u is a r vector valued measurable control, is said to be <u>locally controllable</u> along a solution φ^v corresponding to control v if for some $t_1 > t_o$ all points in some state space (n dimensional) neighborhood of $\varphi^{v}(t_1)$ are attainable in time t_1 by trajectories of (2-9) with admissible control.

It would be somewhat falacious to say that a time dependent system is locally controllable, say at the origin, if all points in a neighborhood of the origin in state space are attainable by trajectories of the system in finite time. To see this, we consider the following example of G. Haynes.

Example 1:

$$\dot{x}_1 = - x_2 + (\cos t)\, u \quad , \qquad x(0) = 0, \ |u(t)| \leq 1$$
$$\dot{x}_2 = x_1 + (\sin t)\, u \ .$$

An integral of the motion is seen to be $x_1 \sin t - x_2 \cos t = 0$, which one can picture as a rotating (with time) line in x_1, x_2 space. As t varies from 0 to 2π, all points of E^2 are swept out by this line. Now multiply the first equation by cos t , the second by sin t and one obtains by adding;

$$\frac{d}{dt}\,(x_1 \cos t + x_2 \sin t) = u \quad \text{or}$$

$$x_1 \cos t + x_2 \sin t = \int_0^t u(\tau)\, d\tau$$. Combining this with the integral of the motion gives

$$x_1^2(t) + x_2^2(t) = \left[\int_0^t u(\tau)\, d\tau\right]^2$$ implying that as time increases, the two dimensional neighborhoods of the origin of E^2 which are attainable also increase.

Since all solutions lie on a surface in (t, x) space, one would hardly feel that the system should be termed locally controllable and is not locally controllable by the definition given above.

We next proceed with an analysis, similar to that used in the papers [9] and [10], to examine local controllability about a given trajectory of the system (2-1). Let $x(t_o) = 0$ be initial data for this system v an arbitrary $\mathscr{L}_2$ control and φ^v the corresponding solution. Let $u(t;\xi)$, $\xi \in E^n$, be a family of controls such that $u(t;0) = v(t)$, u_ξ exists, and denote $x(\cdot\,;\xi)$ as the response to $u(\cdot\,,\xi)$. Then $x(\cdot\,;\xi)$ satisfies

$$x(t;\xi) \equiv \int_{t_o}^{t} \Big[g(\tau, x(\tau;\xi)) + H(\tau, x(\tau;\xi))\, u(\tau;\xi) \Big]\, d\tau .$$

$$x_\xi(t;0) \equiv \int_{t_o}^{t} \Big[g_x(\tau, \varphi^v(\tau)) + H_x(\tau, \varphi^v(\tau)) v(\tau) \Big]\, x_\xi(\tau, 0)$$

$$+\, H(\tau, \varphi^v(\tau))\, u_\xi(\tau, 0)\, d\tau$$

where $H_x v$ is an nxn matrix with i j^{th} element $\sum_{\nu=1}^{r} H^i_{x_j}\, v^\nu$.

For each $t \geq t_o$, we view $x(t;\xi)$ as a mapping $\xi \longrightarrow x$ with $0 \longrightarrow \varphi^v(t)$. Let $Z(t;\varphi^v, u_\xi)$ denote the Jacobian matrix $x_\xi(t;0)$. We have: <u>If for some $\bar{t}$, u_ξ, $Z(t;\varphi^v, u_\xi)$ is non-singular, the attainable set at $\bar{t}$ contains a neighborhood of the point $\varphi^v(\bar{t})$.</u> Let $\Phi(t, t_o)$ be a fundamental solution matrix of the system

$$\dot{x}(t) = \Big[g_x(t, \varphi^v(t)) + H_x(t, \varphi^v(t)) v(t) \Big]\, x(t).$$ Then

$$Z(t; \varphi^v, u_\xi) \equiv \int_{t_o}^{t} \Phi(t,\tau)\, H(\tau, \varphi^v(\tau)) u_\xi\, (\tau; 0)\, d\tau .$$

From lemma I.1 and corollary I.1 we have

<u>Theorem II.2</u> (Kalman) A necessary and sufficient condition that there exist an rxn matrix u_ξ such that $Z(t_1; \varphi^v, u_\xi)$ is non-singular for some $t_1 > t_o$ is that the linear system

$$\dot{y}(t) = \left[g_x(t, \varphi^v(t)) + H_x(t, \varphi^v(t))v(t)\right] y(t) + H(t, \varphi^v(t))u(t)$$

is completely controllable.

In terms of the pfaffian approach the equivalent theorem is

<u>Theorem II.3</u> A necessary and sufficient condition that there exist an rxn matrix u_ξ such that $Z(t_1, \varphi^v, u_\xi)$ is non-singular for some $t_1 > t_o$, is that the pfaffian system $B(t, \varphi^v(t))dx - B(t, \varphi^v(t))$ $\left[g_x(t, \varphi^v(t)) + H_x(t, \varphi^v(t))v(t)\right] x\, dt = 0$ be non-integrable, for some $t_1 \geq t_o$, i.e., that

$$(2\text{-}10) \qquad b(t, \varphi^v(t))dx - b(t, \varphi^v(t)) \left[g_x(t, \varphi^v(t)) + H_x(t, \varphi^v(t))\, v(t))\right] x\, dt$$

is <u>not</u> and exact differential for any b which is a linear combination of the rows of B.

The same method, when applied to a system of the form (2-9) yields

<u>Theorem II.3'</u> A sufficient condition that there exists a $t_1 \geq t_o$ such that all points in some state space neighborhood of $\varphi^v(t_2)$ for all $t_2 > t_1$ are

attainable in time t_2 by trajectories of (2-9) with admissible controls, is that there exists a $t_1 \geq t_o$ such that the pfaffian system

$$B(t; v)dy - B(t; v)\, f_x(t, \varphi^v(t), v(t))y\, dt = 0$$

is not integrable at t_1. [The notation $B(t; v)$ is used to denote the dependence of B on the reference trajectory, specifically $B(t; v)\, f_u(t, \varphi^v(t), v(t)) \equiv 0$.]

It is interesting at this point to see the implications of the assumption that (2-10) <u>is</u> an exact differential. This implies and is implied by

$$\text{(2-11)} \quad \frac{d}{dt} b(t, \varphi^v(t)) \equiv -b(t, \varphi^v(t)) \left[g_x(t, \varphi^v(t) + H_x(t, \varphi^v(t))v(t) \right] ,$$

which can be recognized as the so-called adjoint system of the maximum principle [11] approach to the time optimal problem for system (2-1). It should be noted that if $b(t, \varphi^v(t))$ satisfies (2-11), then it is an adjoint vector which is orthogonal to all of the columns of H. Since the maximum principle (for control components bounded by one in absolute value) implies: choose $u^j(t) = \operatorname{sgn} \sum_{i=1}^{n} b^i(t, \varphi^v(t))H^{ij}(t, \varphi^v(t))$; in this case it yields no information.

I shall designate such a problem as one which admits a <u>totally singular</u> arc φ^v, i.e., where the maximum principle yields no information in the time optimal problem, for any components of the optimal control. The arc would be singular, but not totally singular, if there is an adjoint vector orthogonal to some, but not all columns of H.

<u>Theorem II.4</u> The pfaffian form (2-10) is an exact differential if and only if φ^v is a totally singular arc.

<u>Proof</u>: It has been shown above that if (2-10) is an exact differential, then the vector b satisfies (2-11), which implies φ^v is a totally singular arc.

If φ^v is a totally singular arc, there exists a vector p(t) such that i) $p(t)\, H(t, \varphi^v(t)) \equiv 0$ and ii) $\dot{p}(t) = -p(t)$ $\left[g_x(t, \varphi^v(t)) + H_x(t, \varphi^v(t))\, v(t)\right]$. From i) we conclude that p(t) is a linear combination of the rows of $B(t, \varphi^v(t))$, while II) implies that this linear combination, (2-10), is an exact differential. ∎

To summarize; φ^v <u>not</u> a totally singular arc implies the pfaffian form (2-10) is <u>not</u> an exact differential which implies there exist $\bar{t} \geq t_o$ and $u_{\underline{\underline{\xi}}}$ such that $Z(\bar{t}, \varphi^v, u_{\xi})$ is non-singular and the attainable set at time $\bar{t}$ contains a neighborhood of the point $\varphi^v(\bar{t})$. The contrapositive of this statement provides an interesting characterization of totally singular arcs, i.e., if for every $t_1 > t_o$ there exist points in every state space neighborhood of $\varphi^v(t_1)$ which are not attainable in time t_1 with $\mathscr{L}_2$ controls, the arc φ^v is totally singular. On the other hand, as will be shown by example, a totally singular arc can remain on the boundary of the attainable set, and thus provide a time optimal trajectory.

<u>Theorem II.5</u> If the system (2-1) is not completely controllable at t_o, $Z(t, \varphi^v, u_{\xi})$ is singular for all $t \geq t_o$, u_{ξ} and all reference trajectories φ^v, i.e., every trajectory φ^v is totally singular.

Proof: Any vector b, which is a linear combination of the rows of B, satisfies $b(t, x)H(t, x) \equiv 0$. Thus for any vector $v(t)$,

$\frac{\partial}{\partial x}\left[b(t, x)H(t, x)v(t)\right] \equiv 0$, or $v(t)\, H^T(t, x)b_x(t, x) \equiv$

$-b(t, x)H_x(t, x)v(t)$. Evaluating this identity at the point $(t, \varphi^v(t))$, substituting into (2-11) and expanding of the left side yields

$$\text{(2-12)}\qquad b_t(t, \varphi^v(t)) + b(t, \varphi^v(t))g_x(t, \varphi^v(t)) + g(t, \varphi^v(t))b_x^{\,T}(t, \varphi^v(t)) \equiv v(t)H^T(t, \varphi^v(t)) \left[b_x(t, \varphi^v(t) - b_x^{\,T}(t, \varphi^v(t))\right] .$$

This identity provides a necessary and sufficient condition that (2-10) be an exact differential, i.e., that φ^v be totally singular.

Now assume the system (2-1) is not completely controllable. This means that for some b, a linear combination of the rows of B, the pffafian form $b(t, x)dx - b(t, x)g(t, x)dt$ is an exact differential, or

$$b_t(t, x) \equiv -b(t, x)\, g_x(t, x) - g(t, x)\, b_x^{\,T}(t, x)$$
$$b_x(t, x) \equiv - b_x^{\,T}(t, x) \equiv 0.$$

Evaluating these two identities at $(t, \varphi^v(t))$ for an arbitrary control v shows that (2-12) is satisfied, hence every trajectory φ^v is totally singular.

A conjecture which one might be tempted to make is that if the system (2-1) is completely controllable, it admits no totally singular arcs. This is not true, as the following example from [2] shows.

Example II.1

$$\dot{x}_1 = x_1^2 - x_1^2\, x_2\, u \qquad\qquad x_1(0) = 1$$

$$\dot{x}_2 = -\, x_2 + u \qquad\qquad x_2(0) = 0\ .$$

For the time optimal problem of reaching the point (2, 0), it is shown in [2] that $u \equiv 0$ is the optimal control, if the restriction $|u(t)| \leq 1$ is imposed, and it easily follows that this is also optimal in the class of $\mathscr{L}_2$ controls.

For this problem, one can use for the matrix B, the single vector $b = (1, x_1^2\, x_2)$. The associated pffafian equation is

$$dx_1 + x_1^2\, x_2\, dx_2 + x_1^2\, (x_2^2 - 1)dt = 0.$$

Let $x = (x_1, x_2)$, $a(x) = (1, x_1^2 x_2, x_1^2(x_2^2 - 1))$. Then (curl $a(x)$). $a(x) = 2\, x_2\, x_1^2 \neq 0$, thus the pfaffian is not integrable.

The optimal path from the point (1, 0) to $(\alpha, 0)$, $\alpha > 1$, is obtained with control $u \equiv 0$, and is

$$\varphi^o(t) \quad \begin{cases} \dfrac{1}{1-t} \\ 0\ . \end{cases}$$ This is a totally singular arc. To show this,

we note $b(t, \varphi^o(t)) \equiv (1, 0)$.

$$b(t, \varphi^o(t))dx - b(t, \varphi^o(t)) \left[g_x(t, \varphi^o(t)) + H_x(t, \varphi^o(t)) \circ 0 \right] x\ dt$$

$$= dx_1 + 0\ dx_2 - \frac{2x_1}{1-t}\ dt.$$

Let $\bar{a}(x, t) \equiv (1, 0, \frac{-2x_1}{1-t})$. Then $(\text{curl } \bar{a}) \cdot \bar{a} \equiv 0$ which implies the pfaffian $dx_1 + 0\, dx_2 - \frac{2x_1}{1-t} dt = 0$ is integrable, and φ^0 is a totally singular arc. Here the arc φ^0 is on the boundary of the attainable set.

It should be stressed at this point that it has not been shown that if for some control v, the matrix $Z(t, \varphi^v, u_\xi)$ is singular for all $t \geq t_0$, and u_ξ then sufficiently small n neighborhoods of a point $\varphi^v(t)$ contain points not attainable in time t, from initial data 0 given at t_0. In fact it will next be shown (Example II.2) that this is not the case. To do this we must produce a time optimal problem which posseses a totally singular arc which yields neither a maximum or minimum. Since the arc is totally singular, Theorem II.4 shows that one cannot conclude that the system is locally controllable along this arc by considering the linearized equations as in Theorem II.2. However the use of theorem II.3' on certain arcs which differ from the singular arc but have some points in common with it, will establish the local controllability.

We consider control systems of the form studied in [2], i.e.,

$$(2\text{-}13) \qquad \begin{aligned} \dot{x}_1(t) &= A_1(x(t)) + B_1(x(t))\, u(t) \qquad & x(0) &= x_0 \\ \dot{x}_2(t) &= A_2(x(t)) + B_2(x(t))\, u(t) \qquad & |u(t)| &\leq 1. \end{aligned}$$

We assume that in some region of interest $\mathcal{R}$ of state space,

$$(2\text{-}14) \qquad \Delta(x) \equiv -B_2(x)\, A_1(x) + B_1(x)\, A_2(x) \neq 0$$

and that A_i, B_i, $i = 1, 2$ are C^1 in $\mathcal{R}$.

The pfaffian system associated with (2-13) is the single pfaffian equation

$$(2\text{-}15) \qquad B_2(x)\, dx_1 - B_1(x)\, dx_2 + \Delta(x)\, dt = 0.$$

Since $\Delta(x) \neq 0$ and multiplication by a factor does not change integrability, this can be rewritten as

$$(2\text{-}16) \qquad \frac{B_2(x)}{\Delta(x)}\, dx_1 - \frac{B_2(x)}{\Delta(x)}\, dx_2 + dt = 0.$$

Let $Z(x) = \left(\frac{B_2(x)}{\Delta(x)}, -\frac{B_1(x)}{\Delta(x)}, 1 \right)$; then a necessary and sufficient condition that the pfaffian (2-16) be integrable at a point (t, x) is that $Z(x) \cdot \operatorname{curl} Z(x) \equiv 0$ in a neighborhood of x. Computing yields

$$Z(x) \cdot \operatorname{curl} Z(x) \equiv -\left[\frac{\partial}{\partial x_1} \left(\frac{B_1(x)}{\Delta(x)} \right) + \frac{\partial}{\partial x_2} \left(\frac{B_2(x)}{\Delta(x)} \right) \right] \equiv -\omega(x),$$

where $\omega(x)$ (using the notation of [2]) can be directly computed from the right sides of the differential equations (2-13).

Let v be a continuous control (this is sufficient continuity when the control appears linearly) satisfying $|v(t)| < 1$, and let φ^v be the corresponding trajectory of (2-13).

<u>Theorem II.6</u> If for some $t_1 \geq t_0$, $\varphi^v(t_1)$ is not a zero of ω, then for any $t_2 > t_1$ all points in some state space neighborhood of $\varphi^v(t_2)$ are attainable by trajectories of (2-13), in time t_2, with admissible controls.

Proof: The variational equation for the system (2-13) about the trajectory φ^{v} is given by

$$\dot{y}(t) = \left[A_x(\varphi^{v}(t)) + v(t)\, B_x(\varphi^{v}(t))\right] y(t) + B(\varphi^{v}(t))\, u(t)$$

where $A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$, $B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}$. The pfaffian equivalent to (2-10) for this variational equation is

$$(2\text{-}17)\quad B_2(\varphi^{v}(t))dy_1 - B_1(\varphi^{v}(t))dy_2 + (-B_2(\varphi^{v}(t)),B_1(\varphi^{v}(t)))\left[A_x(\varphi^{v}(t)) + v(t)\, B_x(\varphi^{v}(t))\right] y\, dt = 0.$$

A sufficient condition that (2-17) be <u>not</u> integrable at t_1 is that

$$(2\text{-}18)\quad \frac{d}{dt}(B_2(\varphi^{v}(t)),B_1(\varphi^{v}(t)))\Big|_{t=t_1} \neq (-B_2(\varphi^{v}(t_1)),B_1(\varphi^{v}(t_1)))\left[A_x(\varphi^{v}(t_1)) + v(t_1)\, B_x(\varphi^{v}(t_1))\right],$$

which is implied by $\omega(\varphi^{v}(t)) \neq 0$ as can be shown by a straightforward calculation.

[In terms of Theorem II.4, (2-18) states that $\varphi^{v}(t_1)$ is <u>not</u> a point of a singular arc. In [2, pg. 97] it is shown that for systems of this type singular arcs are characterized by the fact that ω is zero along them. It follows that if $\varphi^{v}(t_1)$ is <u>not</u> a zero of ω, then it is <u>not</u> a point of a singular arc, hence (2-;7) is not integrable and the conclusion of the theorem follows.] ∎

It should be stressed that the integrability of (2-16) requires $\omega(x) = Z(x) \cdot \text{curl } Z(x)$ to be zero in a neighborhood of a point, while Theorem II.6 deals only with the value of ω at a point. It is possible, Example II.1, to have the pfaffian (2-16) not integrable at a point $(\bar{t}, \bar{x})$ at which $\omega(\bar{x}) = 0$, and yet have a trajectory φ^v such that $\varphi^v(\bar{t}) = \bar{x}$ and the system is not locally controllable about φ^v.

We next give the example of a problem which is locally controllable along a totally singular arc.

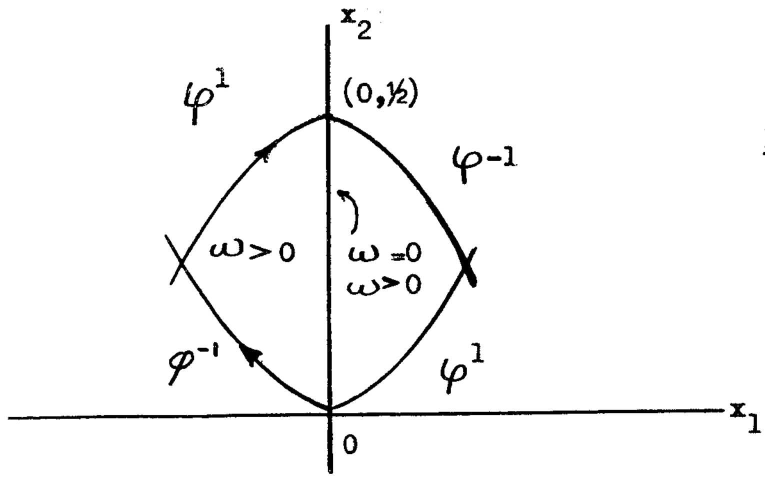
<u>Example II.2</u> (A singular arc $\varphi^0(t)$ such that all points in a neighborhood of $\varphi^0(t_1)$ are attainable in time t_1.)

Consider the system

$$\dot{x}_1 = u \qquad\qquad |u(t)| \leq 1$$

$$\dot{x}_2 = 1 + x_2 x_1^2 u \qquad\qquad x(0) = 0$$

Then $\Delta(x) = 1$, $\omega(x) = x_1^2$, hence if we were to consider the time optimal problem of reaching the final point $x_f(0, ½)$, the Greens theorem approach [2], yields the following



<u>Figure 1</u>

the optimal arc being shown by the arrows. There is an arc along which $\omega = 0$, i.e., $x_1 \equiv 0$, and while this can be attained with the control $u \equiv 0$ it yields neither a maximum or minimum to the time optimal problem. This arc we designate as φ^o;

$$\varphi^o(t) = \begin{cases} \varphi_1{}^o\ (t) \equiv 0 \\ \varphi_2{}^o\ (t) \equiv t. \end{cases}$$

It is easily checked that the variational equation along φ^o is <u>not</u> completely controllable.

Now consider a relation $x_1 = k_1 \sin k_2\, x_2$, $k_1, k_2 > 0$ with $k_2 > 4\pi$. It will be shown that for k_1 sufficiently small, there exists a unique admissible continuous control $\bar{u}(t)$ with trajectory $\varphi^{\bar{u}}$ which has $\left\{(x_1, x_2):\ x_1 = k_1 \sin k_2\, x_2,\ x_2 \geq 0\right\}$ as its track.

From the Greens theorem approach [2] and the symmetry of $\omega(x)$ about the line $x_1 = 0$, the parametrization of $\varphi^{\bar{u}}$ must be such that the even numbered crossings of the x_2 axis, counting only crossings which occur for $x_2 > 0$, one must have

$$\varphi_1^{\bar{u}}\left(\frac{2n\pi}{k_2}\right) = 0 = \varphi_1{}^o\left(\frac{2n\pi}{k_2}\right)$$

$$\varphi_2^{\bar{u}}\left(\frac{2n\pi}{k_2}\right) = \frac{2n\pi}{k_2} = \varphi_2{}^o\left(\frac{2n\pi}{k_2}\right).$$

We will be interested in the case n = 1, so that $\frac{2\pi}{k_2} < 1/2$. It will be shown that there is local controllability along $\varphi^{\bar{u}}$, and since

$\varphi^{\bar{u}}\left(\frac{2\pi}{k_2}\right) = \varphi^{o}\left(\frac{2\pi}{k_2}\right)$, it will follow that a neighborhood of $\varphi^{o}\left(\frac{2\pi}{k_2}\right)$ is attainable in time $\frac{2\pi}{k_2}$.

First we will show that for k_1 sufficiently small, there is a unique continuous u which leads to a trajectory $\varphi^{\bar{u}}$ having $\left\{(x_1, x_2): \ x_1 = k_1 \sin k_2 x_2, \ x_2 \geq 0\right\}$ as its track. Differentiation of the track relation with respect to t yields

$$\dot{x}_1(t) = k_1 k_2 \left[\cos k_2 x_2(t)\right] \dot{x}_2(t).$$

Substitution from the system equations leaves

$$u(t) = k_1 k_2 \left[\cos k_2 x_2(t)\right] \left[1 + x_2(t) x_1^2(t) u(t)\right]. \tag{2-19}$$

For any control u,

$$x_1(t) = \int_0^t u(\tau) d\tau$$

$$x_2(t) = \exp\left[\int_0^t u(\tau) \left(\int_0^t u(\sigma) d\sigma\right)^2 d\tau\right] \int_0^t \exp\left\{-\int_0^{\tau} u(\sigma) \left[\int_0^{\tau} u(\gamma)\, d\gamma\right]^2 d\gamma\right\} d\tau.$$

Substituting these in (2-19) yields an expression of the form

$$u(t) = k_1 (\mathcal{F} u)(t)$$

where the definition of the nonlinear operator $\mathcal{F}$ is obvious. Let

$C\left[0, \frac{1}{2}\right]$ denote the space of continuous vector valued functions u on the interval $\left[0, \frac{1}{2}\right]$, with the supremum norm, and $B^{1/2}$ the closed ball of radius $\frac{1}{2}$ in this space. It is easily shown that for k_1 sufficiently small but positive, $u \in B^{1/2} \Longrightarrow k_1 \mathfrak{F} u \in B^{1/2}$, and $k_1 \mathfrak{F}$ is a contracting map. Thus $k_1 \mathfrak{F}$ has a unique fixed point in $B^{1/2}$, call this point $\bar{u}$. Then $\varphi^{\bar{u}}$ is <u>not</u> a singular trajectory, since k_1 positive implies $\bar{u}(t) \not\equiv 0$, and $\varphi^{\bar{u}}$ has the desired track.

Now for $0 < t_1 < \frac{\pi}{k_2}$, $\varphi^{\bar{u}}(t_1)$ is not a point of the singular arc, hence not a zero of ω. From Theorem II.6 it follows that all points in some neighborhood of $\varphi^{\bar{u}}(t_2)$, for any $t_2 > t_1$ are attainable in time t_2 by trajectories with admissible controls, hence this is true for $t_2 = \frac{2\pi}{k_2}$.

To determine local controllability along $\varphi^{\bar{u}}$ by use of the fundamental solution of the variational equation about this trajectory would be a virtually impossible task.

In concluding, it should be noted that totally singular arcs were defined with no mention made of transversality conditions. It is possible to use these conditions, in very special cases, to rule out the existence of singular arcs in the optimal strategy. Also, for a time optimal problem for a system of the form

$$(2\text{-}20) \qquad \dot{x}(t) = g(x(t)) + H(x(t))u(t)$$

the maximum principle yields the fact that the Hamiltonian is constant along the optimal path. We shall show that this cannot be used to rule out totally singular arcs, since such arcs automatically satisfy the condition even though the Hamiltonian is seemingly a function of time along them.

For the system (2-20) with any given control $u(t)$ we define the Hamiltonian for the time optimal problem as

$$\mathcal{H}(t, x, p) \equiv p \circ g(x) + p \circ H(x)\, u(t) + 1.$$

A necessary condition is that $\mathcal{H}$ is a constant along the optimal trajectory, it need not be so on a non-optimal trajectory. Define the adjoint system as

$$\text{(2-21)} \qquad \dot{p}(t) = -p(t)\, g_x(x,(t)) - p(t)\, H_x(x(t))u(t)$$

Theorem II.7 The Hamiltonian for the system (2-20) is constant along any totally singular arc.

Proof: We defined a totally singular arc as an arc φ^u which satisfies (2-20) for which there exists and adjoint vector $p(t)$ satisfying (2-21) such that $p(t)H(\varphi^u(t)) \equiv 0$ for a set of t values having positive measure. Then

$$\text{(2-22)} \quad \frac{d}{dt}\mathcal{H}(t, \varphi^u(t), p(t)) \equiv \frac{d}{dt}\left[p(t) \cdot g(\varphi^u(t)) + 1\right] \equiv \dot{p}_i g^i + p_i g^i_{x_\nu} \dot{\varphi}^u_\nu \,.$$

From (2-20) $g^i \equiv \dot{\varphi}_i^u - H^{ik} u_k$.

From (2-21) $p^i g^i_{x_\nu} = -\dot{p}_\nu - p_i H^{ik}_{x_\nu} u_k$. Substituting in (2-22)

$$\frac{d}{dt}\mathcal{H}(t,\varphi^u(t),p(t)) \equiv \dot{p}_i\left[\dot{\varphi}_i^u - H^{ik}\, u_k\right] + \left[-\dot{p}_\nu - p_i\, H_x^{ik}\, u_k\right]\dot{\varphi}_\nu^u$$

$$= \left[-\dot{p}_i\, H^{ik} - p_i\, H_{x_\nu}^{ik}\, \dot{\varphi}_\nu^u\right] \quad u_k \equiv -\left\{\frac{d}{dt}\left[p(t)H(\varphi^u(t))\right]\right\} u = 0$$

from the condition $p(t)H(\varphi^u(t)) = 0.$ ∎

REFERENCES TO SECTIONS I AND II

1. Kalman, R. E., Contributions to the Theory of Optimal Control, Boletin De La Sociedad Mathematica Mexicana, 1960, pp. 102-119.

2. Hermes, H., Haynes, G., On the Nonlinear Control Problem with Control Appearing Linearly, SIAM J. on Control, Ser. A., Vol. 1, No. 2, (1963) pp. 85-107.

3. Caratheodory, C., Untersuchungen urber die Grundlagen der Thermodynamik, Mathematische Ann. (1909) pp. 355-386.

4. Chow, W. L., Uber Systeme von Linear Partiellen Differentialgleichungen erster Ordung; Math. Ann. (1940), pp. 95-105.

5. LaSalle, J. P., The Time Optimal Control Problem; Contribution to the Theory of Nonlinear Oscillations, Vol 5, Princeton Univ. Press, 1960, pp. 1-24.

6. Kalman, R. E., Ho, Y. C., Narendra, K. S., Controllability of Linear Dynamical Systems, Contributions to Diff. Eqs., Vol 1, No 2, pp. 189-213.

7. Coddington, E. A., Levinson, N., Theory of Ordinary Differential Equations, McGraw-Hill Book Co., Inc., New York (1955).

8. Milnor, J., Morse Theory, Annals of Math. Studies, No 51, Princeton Univ. Press, Princeton, New Jersey (1963).

9. Lee, E. B., Markus, L., Optimal Control for Nonlinear Processes; Arch. for Rat. Mech. and Anal., Vol 8, No 1, 4.7 (1961), pp. 36-58.

10. Kalman, R. E., Discussion of reference [9], J. Basic Eng., Trans. ASME, Vol 84D (1962).

11. Pontryagin, L. S., Boltyanski, V. G., Gankrelidze, R. V., Mischenko, E. F., The Mathematical Theory of Optimal Processes, Interscience Pub., Div. of J. Wiley and Sons, Inc., New York (1962).

III. THE EQUIVALENCE AND APPROXIMATION OF CONTROL PROBLEMS

INTRODUCTION TO SECTION III

In this section we will be concerned with the time optimal feedback control problem for an n vector system of the form

$$(3\text{-}1) \qquad \mathring{x}(t) = f(t, x(t), u(t)) \quad , \qquad \left(\mathring{x} = \frac{dx(t)}{dt}\right)$$

where the control u is an r vector valued function with values in a given set U. The major interest will be in feedback controls.

One of the difficulties in the theory of optimal feedback control is the discontinuity of the control with respect to the state variables, which the necessary condition termed the maximum principle, so often shows to be the case. Letting $H(t, x, p, u) \equiv p \circ f(t, x, u) - 1$; $u^*(t, x, p)$ be so that $H(t, x, p, u^*(t, x, p)) \geq H(t, x, p, u)$ for all $u \in U$, and $H^*(t, x, p) \equiv H(t, x, p, u^*(t, x, p))$, the Hamilton-Jacobi equation approach [1] often leads to a partial differential equation with discontinuous coefficients, while the Hamiltonian equations of motion which describe the system (the characteristic equations of the Hamilton-Jacobi equation) are of the form

$$(3\text{-}2) \qquad \mathring{x} = \frac{\partial}{\partial p} H^*(t, x, p) \quad , \qquad \mathring{p} = -\frac{\partial}{\partial x} H^*(t, x, p).$$

The maximum principle of Pontriagin, for time optimal problems, assures us that if $u^*(t)$ is an optimal control, $x^*(t)$ the corresponding optimal trajectory, then there exists an absolutely continuous n vector $p^*(t)$, not identically zero, such that $H^*(t,x^*(t),p^*(t)) \equiv H(t,x^*(t),p^*(t),u^*(t))$ while x^* and p^* satisfy equations (3-2). The usual use of the maximum principle proceeds, however, by attempting to generate candidates for an optimal trajectory by solving a two point boundary value problem for the system (3-2). Since u^* may be discontinuous, the fundamental questions of existence and uniqueness of solutions to these equations cannot easily be answered.

An alternative would be to restrict the controls to be continuous, or even C^1, (continuously differentiable) functions and attempt to generate within this class a sequence of controls which will in some sense tend toward the optimal control. In doing this, however, one must seemingly discard the maximum principle which is one of the most useful tools for generating optimal controls, for it so often demands discontinuous controls.

The approach taken here is not to forcefully restrict the class of approximating controls, but instead to generate a class of approximating problems whose solutions will be continuous or C^1 controls and will tend, in a given sense, to the solution of the original problem.

For the system (3-1) let $R(t, x) = \left\{ f(t, x, u) : u \in U \right\}$. We shall say that the time optimal problem for a system $\dot{x} = g(t, x, v)$, $v \in V$ is <u>equivalent</u> to that for the system (3-1) if $\{g(t,x,v) : v \in V\} = R(t, x)$ for all (t,x) in some domain of interest. For given $\epsilon > 0$ we define the time optimal problem for the system $\dot{x} = h^{\epsilon}(t, x, v)$, $v \in V(\epsilon)$ to be an <u>ϵ-approximate equivalent</u> problem to the time optimal problem for (3-1) if $d\left(\left\{h^{\epsilon}(t,x,v) : v \in V(\epsilon)\right\}, R(t, x)\right) < \epsilon$ for all (t,x) in the domain of interest. Here $d(Q,R)$ is the Hausdorff metric distance for sets in E^n.

Intuitively equivalent problems have the same optimal trajectories (as will be shown) while the optimal trajectories of ϵ- approximate equivalent problems will be close (uniformly) to those of the original problem.

It will be shown that under appropriate conditions (essentially the Fillipov existence conditions [2]) the approximating problems can be chosen in such a way that the corresponding feedback controls are continuous, or even of class C^1. In certain cases this allows the Hamilton-Jacobi theory, as derived in [1] , to be utilized for the construction of fields of optimal trajectories and optimal feedback controls.

Although we deal only with the time optimal problem, it should be noted that for a problem of the form $x'(\tau) = f(\tau, x(\tau), u(\tau))$, with the functional to be minimized being $\int_{\tau_0}^{\tau_f} \alpha(\sigma, x(\sigma), u(\sigma))d\sigma$ where the scalar valued function α satisfies $\alpha(\sigma, x, u) \geq \delta > 0$, the change of independent variable

$$t(\tau) = \int_{\tau_0}^{\tau} \alpha(\sigma, x(\sigma), u(\sigma))d\sigma$$ reduces the problem to an

equivalent time optimal problem for the system

$$\dot{y}(t) = \Big[\alpha(\tau(t), y(t), u(t))\Big]^{-1} f(\tau(t), y(t), u(t)) \equiv g(t, y(t), u(t)).$$

THE MAXIMIZATION OF $p \cdot r$ WITH r IN A STRICTLY CONVEX SET

Our motivation is to choose approximating problems for which the maximum principle will yield smooth controls. Let $r^*(p)$ be the function which maximizes the functional $F(p,r) \equiv p \cdot r$ for fixed $p \in E^n - \{0\}$, $r \in R$ a given compact set in E^n. We begin by examining conditions on the set R which will insure that r^* is smooth since it is a maximization of this type which causes discontinuities in the control.

Definition. If S is a set contained in E^n (Euclidean n space) a support hyperplane is a hyperplane M which lies on one side of S and $S \cap M \neq \phi$, the empty set.

Definition. A convex set R contained in E^n will be said to be strictly convex if it contains more than one point, and every support hyperplane has at most one point in common with R.

If R is a compact set in E^n we denote its boundary by ∂R.

Lemma III.1. If R is a strictly convex set in E^n, then R has internal (interior) points. (This result depends on finite dimensionality).

Proof Let r_o, $r_1 \in R$, $r_o \neq r_1$, and V_1 be the linear variety of dimension one determined by these points. Let M_1 be any hyperplane containing V_1. Since M_1 contains two points of R it is not a support plane and there exists a point $r_2 \in R$, $r_2 \notin M_1$. Let V_2 be the linear variety determined by r_o, r_1 and r_2; V_2 has dimension two. Let M_2 be a hyperplane containing V_2. Again there is a point $r_3 \in R$, $r_3 \notin M_2$. We continue inductively getting at the (n-1)st step a linear variety V_{n-1} of dimension (n-1) determined by the points $r_o, \ldots, r_{n-1}$. Then there exists a unique hyperplane M_{n-1} containing V_{n-1}, and again a point $r_n \in R$, $r_n \notin M_{n-1}$. Since R is convex it contains the convex hull of the set of points $r_o, \ldots, r_n$; and since the vectors $r_1 - r_o$, $r_2 - r_o, \ldots, r_n - r_o$ are linearly independent, they determine an n cell which has non void interior. ▮

Lemma III.2. Let R be a strictly convex, compact set in E^n. Then for any fixed $p \in E^n - \{0\}$, the function $F(p, \cdot\,)$ attains its maximum value at a unique point $r^*(p) = r_o \in \partial R$.

Proof For any fixed p, $F(p, \cdot\,)$ is a continuous function on the compact set R and hence attains its maximum there. Suppose the maximum is attained at an interior point $r_o \in R$. Let $N(r_o)$ be a neighborhood of r_o contained in R. Then $p \cdot r_o$ is an interior point of the real interval $p \cdot N(r_o) = \{p \cdot r\colon\ r \in N(r_o)\}$, contradicting the fact that F attains its maximum at r_o.

To show uniqueness, assume $F(p, \cdot\,)$ attains its maximum at r_o, while $r_1 \neq r_o$ belongs to R and $F(p, r_1) = F(p, r_o)$. Define $r(\alpha) = \alpha r_o + (1 - \alpha) r_1$, $-\infty < \alpha < \infty$. It follows that $F(p, r(\alpha)) = F(p, r_o)$ for every such point $r(\alpha)$. If for some α, $r(\alpha)$ is an interior point of R, the argument of the previous paragraph would show a contradiction to $F(p, \cdot\,)$ attaining its maximum at r_o. Thus the one dimensional linear variety $V = \{\alpha r_o + (1-\alpha)r_1\colon -\infty < \alpha < \infty\}$ does not intersect the interior of R, which is not empty by Lemma III.1. By theorem 3.6-E [3] there exists a closed hyperplane M containing V such that the interior of R lies strictly on one side of M. It follows that M is a support plane for R, and since M contains more than one point of R, this is a contradiction to the strict convexity.∎

Theorem III.1 Let R be a strictly convex, compact set in E^n. Then the function $r^*(p)$ (shown to be well defined in lemma III-2) is continuous.

Proof Suppose $p_n \longrightarrow p$. Since R is compact, some subsequence of the sequence $\{r^*(p_n)\}$ converges to a point of R, and there is no loss of generality in assuming it is the original sequence, i.e.

let $r^*(p_n) \longrightarrow r_1$. We suppose $r^*(p) = r_2 \neq r_1$ and seek a contradiction. From the definition of r^*, $F(p, r_2) > F(p, r_1)$; let

$F(p, r_2) - F(p, r_1) = \delta > 0$.

Since F is continuous there exists an $N > 0$ such that

$|F(p_n, r_2) - F(p, r_2)| < \delta/4$ and $|F(p, r_1) - F(p_n, r^*(p_n)| < \delta/4$

for $n \geq N$. Then $F(p_n, r_2) - F(p_n, r^*(p_n)) \equiv \left[F(p, r_2) - F(p, r_1)\right]$

$+ \left[F(p_n, r_2) - F(p, r_2)\right] + \left[F(p, r_1) - F(p_n, r^*(p_n))\right] > \delta/2$ for

$n \geq N$, a contradiction to the definition of $r^*(p_n)$. ∎

We next examine when the function $r^*(p)$ is C^1.

<u>Definition</u>. For $y \in E^n$, $|y| = \left\{ \sum_{1}^{n} y_i^2 \right\}^{1/2}$.

<u>Lemma III.3. Let R be a strictly convex, compact set in E^n which has a unique outward unit normal $n(r)$ at each point $r \in \partial R$. Then for fixed $p \in E^n - \{0\}$, $F(p, \cdot)$ achieves its maximum at the unique point $r_o \in \partial R$ such that $n(r_o) = p/|p|$.</u>

<u>Proof</u> Assume without loss of generality that zero is an interior point of R.

For $x \in E^n$, let $I(x) = \{a: a > 0, a^{-1}x \in R\}$ and define

$\rho(x) = \inf_{a \in I(x)} a.$; $\rho(x)$ is called the support function of R, or also the Minkowski functional. We note that if $r_o \in \partial R$ and y is any vector, then for a real scalar $\alpha > 0$, $\dfrac{\alpha y + r_o}{\rho(\alpha y + r_o)} \in \partial R$

and for α sufficiently small, is in a neighborhood of r_o.

From lemma III.2, we know $F(p, \cdot)$ achieves its maximum at a unique point on ∂R, let r_o be the point. Let $g(y, r_o) = \lim\limits_{\alpha^+ \to 0} \frac{1}{\alpha}\left\{\frac{\alpha y + r_o}{\rho(\alpha y + r_o)} - r_o\right\}$.

Since ∂R has a unique outward normal at each point, $g(y, r_o) = -g(-y, r_o)$. Since $p \cdot r_o \geq p\ r$ for all $r \in \partial R$ in a neighborhood of r_o, it follows that $p \cdot g(y, r_o) \leq 0$ for all y. Assuming there exists y such that $p \cdot g(y, r_o) < 0$ implies $p \cdot g(-y, r_o) > 0$, a contradiction. Thus $p \cdot g(y, r_o) = 0$ for all y, or a necessary condition that r_o presents $F(p, \cdot)$ a maximum is that p be orthogonal to the support hyperplane at r_o. Since R is strictly convex it is easily shown that there are exactly two points which satisfy this necessary condition, one with outward normal $p/|p|$ giving F a maximum, the other with normal $-p/|p|$ which gives F a minimum. ▮

<u>Definition</u>. We say that a strictly convex, compact set R in E^n has a <u>smooth boundary</u> if there exists a unique outward unit normal $n(r) \in C^1$ defined on ∂R. (Actually we consider n as a restriction of a C^1 function in a neighborhood of $r \in \partial R$, see, for example, [4] pg. 27).

<u>Theorem III.2</u>. <u>If R is a compact set in E^n with smooth boundary having positive Gaussian curvature at all points, then $r^*(p) \in C^1$.</u>

<u>Proof</u> Since it is assumed that the unit normal to ∂R is of class C^1, the Gaussian curvature is a continuous positive function on ∂R. But ∂R is compact, thus the Gaussian curvature is bounded away from zero. From theorem 5.5 [5 , pg. 35] it is easily followed that R is strictly convex.

From lemma III.3, we have thar $r^*(p)$ satisfies $n(r^*(p)) = p/|p|$. Let $r_o = r^*(p_o)$ be an arbitrary point on ∂R.

The method will be to utilize the implicit function theorem on a relation of the form $g(r, p) \equiv n(r) - p/|p|$.

Let $\zeta^1, \ldots, \zeta^{n-1}$ be a local coordinate system for a neighborhood of r_o on ∂R. Then the inclusion map from $\partial R \longrightarrow E^n$ determines n smooth functions $x_1(\zeta^1, \ldots, \zeta^{n-1}), \ldots, x_n(\zeta^1, \ldots, \zeta^{n-1})$ or briefly $x(\zeta)$. Assume $x(0) = r_o$ and let V_1 be a measurable neighborhood of zero in the local coordinate system.

Let S^{n-1} be the unit (n-1) sphere; we consider $n(\cdot): \partial R \longrightarrow S^{n-1}$. Define $\theta(\cdot) : V_1 \longrightarrow S^{n-1}$ by $n(x(\zeta)) = \theta(\zeta)$. Thus $n \in C^1 \Longrightarrow \theta \in C^1$.

Let $\varphi = \varphi(p) = p/|p|$, $p \in E^n - \{0\}$; then $\varphi \in C^1$. Our approach will be to utilize the implicit function theorem on the relation $G(\zeta, \varphi) = \theta(\zeta) - \varphi$.

We note that $G \in C^1$, and if $\varphi_o = \varphi(p_o)$ then $G(0, \varphi_o) = 0$. Also $G_\zeta(0, \varphi_o) = \theta_\zeta(0)$. It must be shown that $\det(\theta_\zeta(0)) \neq 0$.

From differential geometry we recall that as ζ varies in V_1, $x(\zeta)$ traces out a region V_2 on ∂R while the normal $\theta(\zeta)$ traces out a region V_3 on the surface of the unit sphere. Let $K(\zeta)$ denote the Gaussian curvature of ∂R at $x(\zeta)$, and A_3 the "area" of V_3. Then

$A_3 = \int_{V_1} K(\zeta) d\zeta$. But $\int_{V_1} \det\left(\frac{\partial \theta(\zeta)}{\partial \zeta}\right) d\zeta = A_3$. Since V_1 is arbitrary (but measurable) and $\theta \in C^1$, this implies $\det\left(\frac{\partial \theta(\zeta)}{\partial \zeta}\right) = K(\zeta)$.

By assumption K is positive at all points of ∂R, hence $\det(\theta_\zeta(0)) \neq 0$. The implicit function theorem now gives the existence of a C^1 function $\zeta(\varphi)$ such that $G(\zeta(\varphi), \varphi) \equiv 0$.

Then $r^*(p) = x(\zeta(\varphi(p))) \in C^1$. ∎

The following is an example of a strictly convex set R with smooth boundary and a point at which the Gaussian curvature K is zero, for which $r^*(p)$ is not C^1.

Let part of the boundary of $R \subset E^2$ consist of the curve $y = x^4$, the rest so as to make R strictly convex and with smooth boundary. We restrict our attention to the defined part of the boundary, in particular to the point (0, 0) at which K is zero.

The outward normal is given by $(4x^3, -1)$. Let $p = (p_1, p_2)$ have p_2 negative and p_1 near zero. To compute $r^*(p) \equiv (x^*(p), y^*(p))$ we compute the point on the curve $y = x^4$ where the normal has direction numbers $(-p_1/p_2, -1)$. This gives $x^*(p) = (-p_1/4 p_2)^{1/3}$, $y^*(p) = (-p_1/4p_2)^{4/3}$, and $\dfrac{\partial x^*(p)}{\partial p_1}$ is seen to not be continuous at $p_1 = 0$.

APPROXIMATION OF OPTIMAL TRAJECTORIES

The Time Optimal Problem

Consider the system (3-1), with U a compact set, and initial data $x(t_o) = x_o$. Let S be a smooth (C^2) manifold in the (n+1) dimensional (t,x) space with the property that for any t_2, t_3, $\{(t,x) \in S: t_2 \leq t \leq t_3\}$ is compact in E^{n+1}. The problem is to find a measurable function u = u(t) having values in U, such that the solution of the initial value problem for (3-1) with u = u(t), intersects the target S in minimum time; i.e., is an optimal trajectory.

We next give the conditions of Fillipov [2], which insure the existence of an optimal (open loop) control, and optimal trajectory for the time optimal problem.

Existence Conditions

(3-3) $f(t,x,u)$ is continuous in all variables t,x and u, and is continuously differentiable with respect to x.

(3-4) $x \cdot f(t,x,u) \leq C(|x|^2 + 1)$ for all t, x, u.

(3-5) $R(t,x) \equiv \{f(t,x,u) : u \in U\}$ is convex for every t,x.

(3-6) There exists at least one measurable function u(t) with values in U, such that the corresponding solution of the initial value, problem for (3-1) attains the target S for some $t_1 \geq t_o$.

Equivalence of Problems

Let the same time optimal problem, as posed for (3-1), also be posed for the system

(3-7) $\dot{x}(t) = g(t,x(t),v(t))$, $v(t) \in V$, a compact set,

where g satisfies condition (3-3). Let $Q(t,x) \equiv \{g(t,x,v) : v \in V\}$.

Theorem III.3 Assume the existence conditions are satisfied for the time optimal problem for the system (3-1). Let $\varphi(\cdot\,; u^*)$ denote the optimal trajectory and u^* the optimal control. Then if $Q(t,x) = R(t,x)$ for all (t,x), $\varphi(\cdot\,; u^*)$ is an optimal trajectory for the time optimal problem for the system (3-7) and there exists a measurable function $v^*(t)$ with values in V such that $\dot{\varphi}(t; u^*) = g(t, \varphi(t;u^*), v^*(t))$ almost everywhere.

Proof $f(t, \varphi(t; u^*), u^*(t))$ is a measurable function of t, with values (almost everywhere) in $R(t, \varphi(t; u^*))$, therefore in $Q(t, \varphi(t; u^*))$. From lemma 1 of Fillipov [2], there exists a measurable function $v^*(t)$ with values in V such that $f(t, \varphi(t;u^*), u^*(t)) = g(t, \varphi(t;u^*), v^*(t))$ almost everywhere. It follows that $\dot{\varphi}(t; u^*) = g(t, \varphi(t;u^*), v^*(t))$ almost everywhere.

Now if $\varphi(\cdot\,; u^*)$ were not an optimal trajectory for (3-7), i.e., $\psi(\cdot\,; v)$ provides a better time, the same argument shows that $\psi(\cdot\,; v)$ is a solution of (3-1) for some measurable control u with values in U, thereby contradicting the assumed optimality of $\varphi(\cdot\,; u^*)$. ∎

This theorem stresses the fact that in seeking optimal trajectories, it is the set function $R(t,x)$ which is of major importance, not the function $f(t,x,u)$ or the control set U.

When the conditions of theorem III.3 are satisfied we define the time optimal problem for the system (3-7) to be equivalent to that for (3-1).

If the existence conditions are satisfied for the time optimal problem, from conditions (3-4) and (3-6) we can obtain a compact region of (t,x) space to which analysis can be restricted. Indeed for $t_o \leq t \leq t_1$ condition (3-4) implies any solution $x(t)$ of (3-1) satisfies $|x(t)|^2 \leq (|x_o|^2 + 1) \exp(2C\,|t_1 - t_o|)$. Here $|x(t)|$ stands for the usual Euclidean norm. Henceforth, we denote by $\mathcal{D}$ the compact region of (t,x)

space defined by $t_o \leq t \leq 2\ t_1$, $|x|^2 \leq (|x_o|^2 + 1)\exp(2C\ |2t_1 - t_o|)$.

<u>Definition</u>. The <u>Hausdorff metric topology</u> for non-empty compact sets in E^n is derived from the following metric: The distance between two non-empty compact sets X and Y in the smallest real number $d = d(X,Y)$ such that X lies in the d neighborhood of Y and Y lies in the d neighborhood of X.

ϵ Approximate Equivalent Problems

<u>Definition</u>. For given $\epsilon > 0$ the time optimal problem for the system $\overset{\circ}{x} = h^\epsilon(t,x,v)$, h^ϵ continuous on $E^1 X E^n X V(\epsilon)$, is said to be an ϵ approximate equivalent problem to the time optimal problem for (3-1) if the set $R(t,x,\epsilon) \equiv \{h^\epsilon(t,x,v) : v \in V(\epsilon)\} \supset R(t,x)$ and $d(R(t,x,\epsilon), R(t,x)) \leq \epsilon$ for all $(t,x) \in \mathcal{A}$.

Since $h^\epsilon(t,x,\cdot)$ is continuous on the compact set $V(\epsilon)$, $R(t,x,\epsilon)$ is compact.

<u>Theorem III.4</u>. Assume that the Fillipov conditions (3-3), (3-4) and (3-5) are satisfied for the time optimal problem with system equations (3-1). Then for every $\epsilon > 0$ there exists an ϵ approximate equivalent problem with system equations $\overset{\circ}{x} = h^\epsilon(t,x,v)$, $v \in V(\epsilon)$ which satisfies the following properties.

a) The control set $V(\epsilon)$ can be taken to be the unit ball of E^n, which we denote B^n.

b) h^ϵ is a C^∞ function on $\mathcal{A} \times B^n$, while for each $(t,x) \in \mathcal{A}$, $h^\epsilon(t,x,\cdot)$ is one-one on $B \longrightarrow E^n$.

c) The set $R(t,x,\epsilon) \equiv \{h^\epsilon(t,x,v) : v \in B^n\}$ has smooth boundary having positive Gaussian curvature.

d) The (single valued) function $v^*(t,x,p)$ with values in B^n which maximizes $H(t,x,p,v;\epsilon) = p\cdot h^{\epsilon}(t,x,v) - 1$ for each $(t,x)\in \mathcal{D}$, $p \in E^n - \{0\}$, is C^1 in t,x, and p. Actually $v^*(t,x,p) \in \partial B^n = S^{n-1}$, the $(n-1)$ sphere.

The proof will proceed by obtaining a simplicial approximation to $\mathcal{D}$ in which the diameters of the simplexes are sufficiently small. For each vertex (t_i, x_i) of a simplex, we approximate the convex set $R(t_i, x_i)$ by a strictly convex set $Q(t_i, x_i, \epsilon)$ having positive Gaussian curvature. A vector function $g^{\epsilon}(t_i, x_i; \cdot)$ is then constructed so that $Q(t_i, x_i; \epsilon) = \{g^{\epsilon}(t_i, x_i; v): v \in B^n\}$, and by use of g^{ϵ}, the set function Q is extended continuously to all of $\mathcal{D}$ in such a manner that for each $(t,x)\in\mathcal{D}$, $Q(t,x;\epsilon)$ has smooth boundary with positive Gaussian curvature. The desired function h^{ϵ} is then obtained by smoothing the function g^{ϵ} in the variables (t,x) via the Friedricks mollifier technique.

<u>Proof</u> $R(t,x)$ is continuous, in the Hausdorff metric topology, on the compact set $\mathcal{D}$. For any $\epsilon > 0$ let $\delta > 0$ be such that $d(R(t,x),R(t',x')) < \epsilon/8$ whenever $|(t,x)-(t',x')| < \delta$. Let σ_g^{n+1} be any bounded geometric simplex which contains $\mathcal{D}$, and K_g be the geometric complex consisting of this single simplex. By barycentric subdivision K_g can be subdivided into a geometric complex K_g' consisting of a family of geometric simplexes $\{\bar{\sigma}_g^{n+1}\}$, each having diameter less than δ.

Each point $(t,x)\in\mathcal{D}$ has a unique representation of the form

$$(t,x) = \sum_{i=1}^{n+2} \alpha_i (t_i, x_i) \quad \text{with } 0 \leq \alpha_i \leq 1, \quad \sum\alpha_i = 1; \text{ where the}$$

$(n+2)$ points (t_i, x_i) are the vertices of the geometric simplex from the family $\{\bar{\sigma}_g^{n+1}\}$ to which the point (t,x) belongs. Without loss of generality we can now consider the union of the members of $\{\bar{\sigma}_g^{n+1}\}$ which have all vertices in $\mathcal{A}$ as a new domain of interest; call this domain again $\mathcal{A}$.

Let (t_i, x_i) be an arbitrary vertex in $\mathcal{A}$. Then $R(t_i, x_i)$ is convex. Let $\mathcal{N}(R(t_i, x_i), \epsilon/4)$ be a convex $\epsilon/4$ neighborhood of $R(t_i, x_i)$. From [6 , pg. 38] there exists a strictly convex set $Q(t_i, x_i, \epsilon)$ containing $\mathcal{N}(R(t_i, x_i), \epsilon/4)$; having an analytic boundary with positive Gaussian curvature, and such that
$d(Q(t_i, x_i, \epsilon), \mathcal{N}(R(t_i, x_i), \epsilon/4)) < \epsilon/4$.
For each $(t_i, x_i) \in \mathcal{A}$ we construct a corresponding set $Q(t_i, x_i, \epsilon)$ as above. We next proceed to define a set valued function $Q(t,x, \epsilon)$ on all of $\mathcal{A}$.

It can be assumed without loss of generality that $0 \in R(t,x)$ for all $(t,x) \in \mathcal{A}$. Indeed if this were not so, one could choose a point $u_o \in U$ and construct new sets $S(t,x) \equiv \{f(t,x,u) - f(t,x,u_o): u \in U\}$ which satisfy this property.

Let B^n be the unit ball in E^n; S^{n-1} its' surface and $v^1, \ldots, v^{n-1}$ a coordinate system on S^{n-1} while v^n measures distance from the origin. Then a ray from the origin through $(v^1, v^2, \ldots, v^{n-1}, 1)$ strikes $\partial Q(t_i, x_i, \epsilon)$ in a unique point which we denote $g^{\epsilon}(t_i, x_i, v^1, \ldots, v^{n-1}, 1)$. This defines $g^{\epsilon}(t_i, x_i, \cdot)$ on S^{n-1}; to extend it to B^n let $v = (v^1, \ldots, v^n)$ B^n. Define $g^{\epsilon}(t_i, x_i, v)$ as that point in $Q(t_i, x_i, \epsilon)$ which lies on the ray through the origin and $(v^1, \ldots, v^{n-1}, 1)$ and is such that

$$\frac{|g^{\epsilon}(t,x,v)|}{|g^{\epsilon}(t,x,v^1,\ldots,v^{n-1},1|} = v^n.$$

Then $g^{\epsilon}(t_i, x_i, \cdot): B^n \longrightarrow Q(t_i, x_i, \epsilon)$ in a one to one fashion. We will define $Q(t, x, \epsilon)$ on all of Δ by extending the definition of g^{ϵ} to all $(t,x) \in \Delta$.

Assume $(t,x) \in \Delta$. Let $(t,x) = \sum_{i=1}^{n+2} \alpha_i (t_i, x_i)$ be the unique representation of (t,x) in terms of the vertices of the geometric simplex of K'_g to which it belongs. Define

$g^{\epsilon}(t,x,v) = \sum_{i=1}^{n+2} \alpha_i \, g^{\epsilon}(t_i, x_i, v)$, $v \in B^n$. Then if

$Q(t,x,\epsilon) = \left\{ g^{\epsilon}(t,x,v) : v \in B^n \right\}$ it follows that:

i) $\eta(R(t,x), \epsilon/8) \subset Q(t,x,\epsilon)$. Indeed, from the choice of δ, $\eta(R(t,x), \epsilon/8) \subset \eta(R(t_i, x_i), \epsilon/4) \subset Q(t_i, x_i, \epsilon)$ for all vertices (t_i, x_i) of the simplex in which (t,x) is contained. But $Q(t,x,\epsilon) = \sum \alpha_i Q(t_i, x_i, \epsilon)$. Thus if a point is in $\eta(R(t, x), \epsilon/8)$ it is in $Q(t,x,\epsilon)$.

ii) $d(Q(t,x,\epsilon), R(t, x)) \leq 3\epsilon/4$. To show this one notes that $R(t_i, x_i) \subset \eta(R(t_j, x_j), \epsilon/4) \subset Q(t_j, x_j, \epsilon)$ for all $i, j = 1, 2, \ldots, n+2$. Therefore

$$
\begin{aligned}
d(R(t,x), Q(t,x,\epsilon)) &\leq d(R(t,x), R(t_i, x_i)) + \\
d(R(t_i, x_i), \sum_j \alpha_j Q(t_j, x_j, \epsilon)) &\leq \epsilon/8 + \\
\max_j \left[d(R(t_i, x_i), Q(t_j, x_j, \epsilon) \right] &\leq \epsilon/8 + \\
\max_j \left[d(R(t_i, x_i), R(t_j, x_j)) + d(R(t_j, x_j), Q(t_j, x_j, \epsilon)) \right] &\leq 3\epsilon/4.
\end{aligned}
$$

iii) $Q(t,x,\epsilon)$ is strictly convex, with smooth boundary having positive Gaussian curvature, for each (t,x) . Indeed of $K(t,x,v^1,\dots,v^{n-1})$ is Gaussian curvature at the point $g(t,x,v^1,\dots,v^{n-1}, 1) \in \partial R(t,x,\epsilon)$, then $K(t,x,v^1,\dots,v^{n-1}) = \sum_{i=1}^{n+2} \alpha_i K(t_i, x_i, v^1,\dots,v^{n-1})$.

iv) From the construction, $g^\epsilon(t,x,v)$ is analytic in v for fixed (t,x) and continuous in (t,x) for fixed v.

Combining the results of i) and ii) shows that for $(t,x) \in \mathcal{D}$,

$\mathcal{N}(R(t,x), \epsilon/8) \subset Q(t,x,\epsilon) \subset \mathcal{N}(R(t,x), 3\epsilon/4)$.

It will next be shown that using $g^\epsilon(t,x,v)$ one can construct a mapping $h^\epsilon(t,x,v)$ on $\mathcal{D} \times B^n \longrightarrow E^n$ such that if $R(t,x,\epsilon) = \left\{ h^\epsilon(t,x,v) : v \in B^n \right\}$, then $R(t,x,\epsilon)$ is a strictly convex, compact set containing $R(t,x)$; $d(R(t,x,\epsilon), R(t,x,)) < \epsilon$; $\partial R(t,x,\epsilon)$ is smooth with positive Gaussian curvature, and if $n(t,x,h^\epsilon(t,x,v^1,\dots,v^{n-1},1)$ is a unit normal to $\partial R(t,x,\epsilon)$ at $h^\epsilon(t,x, v^1,\dots,v^{n-1}, 1)$ then it is a C^1 function of all arguments.

For simplicity of notation let $y = (t,x)$ denote a point in $\mathcal{D}$, and let $S^k(y-\bar{y})$ be a mollifier function; see [7] . As an example one could choose $S^k(y-\bar{y}) = (k/4\pi)^{\frac{n+1}{2}} \exp\left\{ -\frac{k}{4} \left[\sum_{i=1}^{n+1} (y^i - \bar{y}^i)^2 \right] \right\}$.

Extend $g^\epsilon(y,v)$ as the zero function for y in the complement of $\mathcal{D}$.

Define $h^k(y,v) = \int_{E^{n+1}} S^k(y-\bar{y})\, g^\epsilon(\bar{y},v)\, d\bar{y}$.

Then for every integer $k \geq 0$, h^k is an analytic function, while h^k and its derivatives with respect to v tend uniformly to g^ϵ and its derivatives with respect to v.

Let $R^k(t,x,\epsilon) \equiv \left\{h^k(t,x,v): v \in B^n\right\}$. Since the Gaussian curvature to $\partial Q(t,x,\epsilon)$ is given as a multilinear combination of the derivatives $g^\epsilon_{v^i}(t,x,v^1,\ldots,v^{n-1},1)$ while the curvature of $\partial R^k(t,x,\epsilon)$ is given by the same multilinear combination of the derivatives $h^k_{v^i}(t,x,v^1,\ldots,v^{n-1},1)$; one can choose k sufficiently large so that $\partial H^k(t,x,\epsilon)$ has positive Gaussian curvature while $R(t,x) \subset H^k(t,x,\epsilon) \subset \mathcal{N}(R(t,x),\epsilon)$. For such a choice of k, define $h^\epsilon(t,x,v) = h^k(t,x,v)$, $R(t,x,\epsilon) \equiv \left\{h^\epsilon(t,x,v): v \in B^n\right\}$.

From its construction, h^ϵ satisfies conclusions a), b) and c), while a unit normal $n(t,x,h^\epsilon(t,x,v^1,\ldots,v^{n-1},1))$ to $\partial R(t,x,\epsilon)$ is a C^1 function of $(t,x, v^1,\ldots,v^{n-1})$.

It remains to show part d). Using lemma III.3 define $r^*(t,x,p;\epsilon)$ as the unique point on $\partial R(t,x,\epsilon)$ such that $n(t,x,r^*(t,x,p,\epsilon)) = p/|p|$. It will be shown that r^* is a C^1 function of t, x, and p by a proof similar to that of theorem III.2. Defining $v^*(t,x,p)$ as the unique point on ∂B^n such that $h^\epsilon(t,x,v^*(t,x,p)) = r^*(t,x,p,\epsilon)$ it follows that v^* maximizes $H(t,x,p,v;\epsilon)$ and it will be shown that v^* is a C^1 in t, x and p.

For fixed (t,x), we have

$$S^{n-1} \xrightarrow{h^\epsilon(t,x,v^1,\ldots,v^{n-1},1)} \partial R(t,x,\epsilon) \xrightarrow{n(t,x,r)} S^{n-1}$$

which naturally induces a map $\Theta(t,x,v^1,\ldots,v^{n-1})$ from $S^{n-1} \longleftrightarrow S^{n-1}$ defined by $\Theta(t,x,v^1,\ldots,v^{n-1}) \equiv n(t,x,h^\epsilon(t,x,v^1,\ldots,v^{n-1},1))$. Since we are only interested in $\partial B^n = S^{n-1}$, no confusion should occur if for the remainder of this argument we let $v = (v^1,\ldots,v^{n-1}) \in S^{n-1}$ and therefore write $\Theta(t,x,v)$. This will be done.

Let $\varphi = \varphi(p) = p/|p|$, $p \in E^n - \{0\}$ and define $G(t,x,v,\varphi) \equiv \Theta(t, x, v,) - \varphi$. We will apply the implicit function theorem to G, which is easily seen to be a C^1 function. For each t_o, x_o, $\varphi_o = p_o/|p_o|$, there exists a unique point $r_o = r^*(t_o, x_o, p_o; \epsilon)$ such that if $n(t_o, x_o, r_o) = p_o/|p_o|$ and v_o is the unique point on S^{n-1} such that $h^\epsilon(t_o, x_o, v_o) = r_o$, then $G(t_o, x_o, v_o, \varphi_o) = 0$. One next notes that $G_v(t_o, x_o, v_o, \varphi_o) = \Theta_v(t_o, x_o, v_o)$, and from the definition of Θ (see also the proof of theorem III.2) $\det\left[\Theta_v(t_o, x_o, v_o)\right]$ is the Gaussian curvature at $r_o \in \partial R(t, x, \epsilon)$ which is positive. The implicit function theorem yields the existence of a C^1 function $v(t, x, \varphi)$ such that $G(t, x, v,(t, x, \varphi), \varphi) \equiv 0$ in a neighborhood of the arbitrary point t_o, x_o, φ_o. Then $r^*(t, x, p; \epsilon) \equiv h^\epsilon(t, x, v(t, x, \varphi(p))) \in C^1$, while $v^*(t, x, p) \equiv v(t, x, \varphi(p))$ is also C^1. ∎

THE RELATION OF TRAJECTORIES OF THE APPROXIMATING PROBLEM TO THOSE OF THE TIME OPTIMAL PROBLEM

We assume the system (3-1) satisfies the Fillipov existence conditions (3-3), (3-4), (3-5) and (3-6), with t_1 a time in which the target set S is attainable. For any $\epsilon > 0$ let $h^{\epsilon}(t, x, v)$, $v \in V(\epsilon)$, be an ϵ approximate equivalent problem (not necessarily having the special properties shown to exist in theorem III.4). From condition (3-6) and the relation $R(t, x, \epsilon) \supset R(t, x)$, it readily follows that for every $\epsilon > 0$ there exists at least one measurable function v with values in $V(\epsilon)$ such that the corresponding trajectory $\varphi^{\epsilon}(\cdot\, ; v)$ of the ϵ approximate problem attains the target S.

It will next be shown that when dealing with the approximate problem, analysis can again be restricted to a compact set. Indeed any vector $h^{\epsilon}(t, x, v)$ can be written as $f(t, x, u) + \alpha(t, x)$ where $|\alpha(t, x)| \leq \epsilon$. Then for any trajectory x(t) of the approximate problem

$$\frac{1}{2}\frac{d}{dt}|x(t)|^2 = x(t)\cdot h^{\epsilon}(t,x(t),v(t)) = x(t)\cdot f(t,x(t);u(t)) + x(t)\cdot \alpha(t,x(t))$$

$$\leq C(1 + |x(t)|^2) + \epsilon |x(t)| .$$

$$\frac{d}{dt} \ln (1 + |x(t)|^2) \leq 2C + \frac{2\epsilon|x(t)|}{1 + |x(t)|^2} \leq 2(C + \epsilon),$$

$$|x(t)|^2 \leq (1 + |x_o|^2)e^{2(C+\epsilon)(t_1 - t_o)} .$$

Define $\mathcal{A}^{\epsilon}$ to be the compact region in E^{n+1} dimensional (t,x) space so that $|x|^2 \leq (1 + |x_o|^2) \exp\left[2(C+\epsilon)(2t_1 - t_o)\right]$, $t_o \leq t \leq 2t_1$.

Theorem III.5. Consider a sequence $\{\epsilon_k\}$ with $\epsilon_k > 0$, $\epsilon_k \longrightarrow 0$ and let φ^{ϵ_k} denote the time optimal trajectory (assumed to exist) for the ϵ_k approximate problem. Then $\{\varphi^{\epsilon_k}\}$ is an equicontinous family on the interval $[t_o, t_1]$. It has a uniformly convergent subsequence which converges to a function φ^* having the following properties.

i) φ^* is absolutely continuous

ii) There exists a measurable function u^* with values in U such that $\dot{\varphi}^*(t) = f(t, \varphi^*(t), u^*(t))$ almost everywhere.

iii) There exists a smallest $t^* \geq t_o$ such that $\varphi^*(t^*) \in S$

iv) φ^* is a time optimal trajectory for the system (3-1).

Proof We shall prove the conclusions in the order that they are stated.

Without loss of generality, assume that $R(t,x,\epsilon_1) \supset R(t,x,\epsilon_2) \supset \ldots \supset R(t,x)$. Therefore analysis can be restricted to the compact region $\mathcal{A}^{\epsilon_1}$. Our first goal is to show that there is a constant N independent of ϵ_k such that φ^{ϵ_k} is Lipschitz continuous with Lipschitz constant N. To accomplish this, for a compact set R in E^n let $\rho(R)$ denote $\max_{r \in R} |r|$. For fixed ϵ_1, $R(t, x, \epsilon_1)$ is a continuous set valued function (in the Hausdorff metric topology) on the compact set $\mathcal{A}^{\epsilon_1}$ and therefore the composite map $\rho(R(t, x, \epsilon_1))$ is a continuous real valued function on $\mathcal{A}^{\epsilon_1}$, hence bounded. Let N be its bound. It follows that $|h^{\epsilon_k}(t, x, v)| \leq N$ for all ϵ_k and any trajectory φ^{ϵ_k} is Lipschitz continuous with Lipschitz constant N. Thus $\{\varphi^{\epsilon_k}\}$ is equicontinuous and has a subsequence which converges uniformly to a Lipschitz continuous function φ^*, which is therefore absolutely continuous. We will not distinguish between $\{\varphi^{\epsilon_k}\}$ and its convergent subsequence.

ii) We next show that for almost all $t \in [t_o, t_1]$, $\dot{\varphi}^*(t) \in R(t, \varphi^*(t))$.

Since the set function $R(t, x)$ is continuous in the Hausdorff metric topology (a consequence of the continuity of f), for any $\nu > 0$ let $R_\nu(t, x)$ be a closed convex ν - neighborhood of $R(t, x)$. Then $R_\nu(t, x)$ is also a continuous set function.

Since $\dot{\varphi}^{\epsilon_k}(t) \in R(t, \varphi^{\epsilon_k}(t), \epsilon_k)$ and $R(t, x, \epsilon_k) \longrightarrow R(t, x)$ in the Hausdorff metric topology, there exists and N such that for all $n \geq N$, $\dot{\varphi}^{\epsilon_k}(t) \in R_\nu(t, \varphi^*(t))$. Fillipov's proof of theorem 1, [2] now applies to show that for almost all t, $\dot{\varphi}^*(t) \in R_\nu(t, \varphi^*(t))$. But $R(t, x)$ is closed and ν arbitrarily small, hence $\dot{\varphi}^*(t) \in R(t, \varphi^*(t))$ for almost all t.

From the lemma of Fillipov [2], we then obtain the existence of a measurable control u* with values in U, such that for almost all $t \in [t_o, t_1]$, $\dot{\varphi}^*(t) = f(t, \varphi^*(t), u^*(t))$.

iii) Let $t_{\epsilon_k} > t_o$ denote the optimal time for the ϵ_k approximate problem. Since $R(t, x, \epsilon_1) \supset R(t, x, \epsilon_2) \supset \dots$ it follows that $\{t_{\epsilon_k}\}$ is a monotone non-drecreasing sequence of reals bounded above by t_1. Let t* be its limit. Now $\varphi^{\epsilon_k}(t_{\epsilon_k}) \in S$ for each k, and $\{(t,x) \in S: t_o \leq t \leq t_1\}$ is compact in E^{n+1}, thus $\varphi^{\epsilon_k}(t_{\epsilon_k}) \longrightarrow \varphi^*(t^*) \in S$.

iv) Suppose φ^* is not a time optimal trajectory for the system (3-1). Then there exists a measurable control u with values in U and corresponding trajectory $\varphi(\cdot\,; u)$ such that $\varphi(t_o; u) = x_o$, $\varphi(t_3; u) \in S$ and $t_3 < t^*$. This implies that for k sufficiently large, $t_3 < t_{\epsilon_k}$; but $\varphi(\cdot\,; u)$ is an admissible trajectory to all ϵ approximate problems. This contradicts the optimality of φ^{ϵ_k}. ∎

This theorem essentially tells us that for sufficiently small ϵ, the optimal trajectories of the ϵ approximate problem are uniformly close to an optimal trajectory of the original problem.

In the next section the "smoothness" which theorem III.4 shows is possible for the feedback control of the ϵ approximate problem, will be exploited to obtain solutions.

Hamilton-Jacobi Theory

Let the time optimal problem for (3-1) satisfy the Fillipov existence conditions. Let $\dot{x} = h^{\epsilon}(t, x, v)$ denote an ϵ approximate system with the properties a), b), c) and d), shown to exist in theorem III.4. For the time optimal problem associated with the approximate problem we define the functions

$$H(t, x, p, v, \epsilon) \equiv p \cdot h^{\epsilon}(t, x, v) - 1$$

$$H^*(t, x, p, \epsilon) \equiv H(t, x, p, v^*(t, x, p), \epsilon).$$

The inequality

$$(3\text{-}8) \qquad H(t, x, p, v^*, \epsilon) > H(t, x, p, v, \epsilon) \text{ for all } v \in B^n,\ v \neq v^*$$

is a consequence of the definition of v^*.

For the sake of completeness we repeat a short argument of Kalman ([1], pp. 321-322) to show that for fixed $\epsilon > 0$,

$$H^*_x(t, x, p, \epsilon) \equiv p \cdot h^{\epsilon}_x(t, x, v^*(t, x, p))$$

$$H^*_p(t, x, p, \epsilon) \equiv h^{\epsilon}(t, x, v^*(t, x, p)).$$

Indeed, we know that $v^*(t, x, p) \in \partial B^n = S^{n-1}$, thus let $g(v)$ be a smooth relation such that $g(v) = 0$ determines S^{n-1} in a neighborhood of $v^*(t, x, p)$. Then $g_v(v^*(t, x, p))\, v^*_x(t, x, p) \equiv 0$ and $g_v(v^*(t, x, p))\, v^*_p(t, x, p) \equiv 0$.

Noting that v^* maximizes $H(t, x, p, v, \epsilon)$, we consider this maximization subject to the constraint $v \in S^{n-1}$, i.e., $g(v) \equiv 0$. The Lagrange multiplier rule implies $H_v + \nu g_v = 0$ where $\nu \neq 0$. Evaluting this at v^* and multiplying on the right by $v_x^*(t, x, p)$ and $v_p^*(t, x, p)$, in turn, gives the required result.

If φ^ϵ, ψ^ϵ are solutions, respectively, to the boundary value problem

$$(3\text{-}9) \qquad \dot{x} = H_p^*(t, x, p, \epsilon) \equiv h^\epsilon(t, x, v^*(t, x, p))$$

$$(3\text{-}10) \qquad \dot{p} = -H_x^*(t, x, p, \epsilon) \equiv - p \cdot h_x^\epsilon (t, x, v^*(t, x, p))$$

with boundary data $x(t_o) = x_o$, $x(t_1) = x_1$, then (3-8) shows that $v^*(t, \varphi^\epsilon(t), \psi^\epsilon(t))$ satisfies the necessary condition termed the maximum principle, for being an optimal (open loop) control for the time optimal problem of attaining the state x_1 from the state x_o for the approximating system.

It should be noted that under the conditions assumed, $v^* \in C^1$ and the initial value problem for the equations (3-9), (3-10) with data given at t_o will have a unique solution in a neighborhood of t_o. If v^* is discontinuous, this presents a serious difficulty in the application of the maximum principle.

With the (Hamiltonian) function $H^*(t, x, p, \epsilon)$, $\epsilon > 0$ and fixed, we associate the Hamilton-Jacobi partial differential equation

$$(3\text{-}11) \qquad V_t(t,x) + H^*(t, x, V_x(t, x), \epsilon) = 0.$$

Let the target S be a "smooth" n-dimensional, non-characteristic manifold in the (n+1) dimensional (t,x) space, and prescribe the Cauchy data $V(t,x) = 0$, $(t,x) \in S$. The solution, in the classical sense, of this partial differential equation problem, we denote by V^ϵ ; the domain of solution by $\mathcal{D}(\epsilon, S)$.

The characteristic equations associated with (3-11) are the equations (3-9), (3-10). If a point (t_o, x_o) is in $\mathcal{A}(\epsilon, S)$ there exists a point $(t_1, x_1) \in S$ such that the boundary value problem consisting of the equations (3-9), (3-10) with boundary data for (3-10) being $x(t_o) = x_o$, $x(t_1) = x_1$, has a solution. The solution of such a boundary value problem, when it exists, will be denoted by φ^ϵ, ψ^ϵ.

From the continuity condition, for each $\epsilon > 0$, $V_x^\epsilon(t, \varphi^\epsilon(t))$ exists and satisfies equation (3-10). (See for example [1]). Thus we can make the association $\psi^\epsilon(t) \equiv V_x^\epsilon(t, \varphi^\epsilon(t))$.

Let $\mathcal{A}^-(\epsilon, S)$ denote the set of points $(t_o, x_o) \in \mathcal{A}(\epsilon, S)$ for which $t_o \leq t_1$; (t_1, x_1) being the point on S joined to (t_o, x_o) by a curve φ^ϵ. Assume $(t_o, x_o) \in \mathcal{A}^-(\epsilon, S)$. If we use the initial data $x(t_o) = x_o$, $p(t_o) = V_x^\epsilon(t_o, x_o)$; by virtue of knowing a solution of the partial differential equation we have the proper initial data to reduce the previous two point boundary value problem for (3-9) and (3-10) to an initial value problem. Thus to determine the trajectory φ^ϵ we can consider the system

$$\text{(3-12)} \qquad \dot{x} = H_p^*(t, x, V_x^\epsilon(t, x); \epsilon), \qquad x(t_o) = x_o.$$

The major advantage of this method is that now $v^* = v^*(t, x, V_x^\epsilon(t,x))$, i.e., a feedback control.

<u>Theorem III.6</u> (Kalman) Assume $(t_o, x_o) \in \mathcal{A}^-(\epsilon, S)$; V^ϵ is the solution of the Hamilton-Jacobi equation (3-11) and φ^ϵ the solution of (3-12). Then φ^ϵ is a time optimal trajectory relative to all trajectories $\varphi(\cdot\,; v)$ which connect (t_o, x_o) to S and lie in $\mathcal{A}^-(\epsilon, S)$.

<u>Proof</u> Assume, without loss of generality, that $(t_o, x_o) \notin S$. From the definition of H, H* and V^ϵ,

$$0 \equiv V_t^\epsilon(t, x) + V_x^\epsilon(t, x) \cdot h^\epsilon(t, x, v^*(t, x, V_x^\epsilon(t, x))) - 1 > V_t^\epsilon(t, x) +$$

$$V_x^\epsilon(t, x) \cdot h^\epsilon(t, x, v) - 1 \qquad \text{for all } v \in B^n,\ v \neq v^*.$$

Assume that t_ϵ $(t_\epsilon > t_o)$ is the first time such that $(t_\epsilon, \varphi^\epsilon(t_\epsilon)) \in S$. Let Ω^ϵ denote the set of measurable control functions having values in B^n and leading to trajectories of the ϵ approximate problem which connect (t_o, x_o) with a point on S and lie in $\mathcal{A}^-(\epsilon, S)$. Then Ω^ϵ is not empty since $(t_o, x_o) \in \mathcal{A}^-(\epsilon, S)$ and φ^ϵ a characteristic implies $\{(t, \varphi^\epsilon(t)): t_o \leq t \leq t_\epsilon\}$ is in $\mathcal{A}^-(\epsilon, S)$. If $v^*(t, \varphi^\epsilon(t), V_x^\epsilon(t, \varphi^\epsilon(t))$ is the only function (to within a set of zero measure) in Ω^ϵ, the result is trivially true. If this is not the case let $v = v(t)$ be any function in Ω^ϵ differing from $v^*(t, \varphi^\epsilon(t), V_x^\epsilon(t, \varphi^\epsilon(t)))$ on a set Λ of positive measure. Let $\varphi(\cdot\,; v)$ be the corresponding solution of the approximate system and t_2 the first time such that $(t_2, \varphi(t_2; v)) \in S$. $(t_2 > t_o)$. We must show $t_\epsilon < t_2$.

Calculating

$$\frac{d}{dt} V^\epsilon(t, \varphi(t; v)) - 1 \equiv V_t^\epsilon(t, \varphi(t;v)) + V_x^\epsilon(t, \varphi(t;v)) \cdot h^\epsilon(t, \varphi(t;v), v(t)) - 1 \leq 0$$

for all t and strictly less than zero for $t \in \Lambda$, implying $V^\epsilon(t_2, \varphi(t_2; v)) - V^\epsilon(t_o, x_o) < t_2 - t_o$. But $V^\epsilon(t_2\ \varphi(t_2; v)) = 0$ since $(t_2, \varphi(t_2; v)) \in S$, yielding $-V^\epsilon(t_o, x_o) < t_2 - t_o$. Similarly $\frac{d}{dt} V^\epsilon(t, \varphi^\epsilon(t)) - 1 \equiv 0$ implying $-V^\epsilon(t_o, x_o) = t_\epsilon - t_o$. Combining the last two inequalities gives $t_\epsilon <- t_2$ as was to be shown. ∎

THE CONSTRUCTION OF APPROXIMATING PROBLEMS WHEN THE CONTROL APPEARS LINEARLY.

Theorem III.4 gives conditions for the existence of an ϵ equivalent approximate problem which has the unit ball B^n as the set of values which the control can assume. However, the functional form of the approximating system is allowed to vary with ϵ.

In this section we consider a system of the form

$$\dot{x}(t) = g(t, x(t)) + H(t, x(t))\, u(t), \tag{3-13}$$

$u(t) \in U$, a compact convex set in E^r with $1 \leq r \leq n$; H an $n \times r$ matrix valued C^2 function; while g is a C^2, n vector valued function. For such systems it is possible to provide a simple construction for ϵ approximate problems.

Since, for the approximate problem, one desires $R(t, x, \epsilon)$ to be strictly convex and lemma III.1 shows this implies non void interior, one is led to extend H to an $n \times n$ matrix valued function and approximate the control set by a compact set $V(\epsilon)$ which contains U. Furthermore, $V(\epsilon)$ should have a non-void n dimensional interior, a smooth boundary with positive Gaussian curvature, and be such that in the Hausdorff metric topology, $\lim_{\epsilon \to 0} V(\epsilon) = U$.

The method of construction and the application to approximating problems will be demonstrated in a two dimensional example; its generalization to higher dimensions being immediate.

Example III.1 (Bushaw control problem).

Consider the time optimal problem for the system

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1 + u \end{aligned} \tag{3-14}$$

with arbitrary initial data $x(0) = x_o$, and target $S = \{(t, x_1, x_2): x_1 = 0, x_2 = 0\}$. The control u is to satisfy $-1 \leq u(t) \leq 1$, i.e., $U = [-1,1]$.

As an ϵ approximate problem we take the system

$$(3\text{-}15) \qquad \begin{aligned} \dot{x}_1 &= x_2 + v_1 \\ \dot{x}_2 &= -x_1 + v_2 \end{aligned}$$

with the same initial data and target, but with $V(\epsilon) = \left\{ v \in E^2 \colon v_1^{\,2} + \epsilon^2 v_2^{\,2} \leq \epsilon^2 \right\}$, i.e., an ellipse with semi major axis 1 and semi minor axis ϵ. Thus in the Hausdorff metric topology $\lim_{\epsilon \to 0} V(\epsilon) = U$, and $\partial R(t, x, \epsilon)$ is smooth with positive Gaussian curvature. From the Hamilton-Jacobi theory

$$H(t, x, p, v, \epsilon) \equiv p_1 x_2 + p_1 v_1 - p_2 x_1 + p_2 v_2 - 1.$$

Using lemma III.3 one computes

$$v^*(t, x, p^* \equiv (\epsilon^2 p_1 \left[\epsilon^2 p_1^{\,2} + p_2^{\,2}\right]^{-\frac{1}{2}}, \; p_2 \left[\epsilon^2 p_1^{\,2} + p_2^{\,2}\right]^{-\frac{1}{2}}$$

from which it follows that

$$H^*(t, x, p, \epsilon) \equiv p_1 x_2 - p_2 x_1 + \left[p_1^{\,2} \epsilon^2 + p_2^{\,2}\right]^{-\frac{1}{2}} - 1.$$

The associated Hamilton-Jacobi equation is

$$(3\text{-}16) \quad V_t(t,x) + x_2 V_{x_1}(t,x) - x_1 V_{x_2}(t,x) + \left[\epsilon^2 V_{x_1}^{\,2}(t,x) + V_{x_2}^{\,2}(t,x)\right]^{\frac{1}{2}} - 1 = 0.$$

Since the independent variables appear linearly, while the dependent variable has derivatives which appear non-linearly, the Legendre contact transformation is suggested. Let $V(t,x) = W(t,p) - p \cdot x$. Then $V_t = W_t$, $V_x = -p$, $W_p = x$ and the transformed equation is

$$W_t(t,p) - p_1 W_{p_2}(t,p) + p_2 W_{p_1}(t,p) + \left[\epsilon^2 p_1^{\,2} + p_2^{\,2}\right]^{\frac{1}{2}} - 1 = 0.$$

The characteristic equations associated with this linear partial differential equation are $t'(\tau) = 1$, $p_1'(\tau) = p_2(\tau)$, $p_2'(\tau) = -p_1(\tau)$, yielding solutions: $t = \gamma + \tau$, $p_1 = \alpha \sin(\tau + \beta)$, $p_2 = \alpha \cos(\tau + \beta)$ with α, β, γ arbitrary constants. Then $\frac{d}{dt} W(t(\tau), p(\tau)) = 1 - \left[\epsilon^2 p_1^{\,2}(\tau) + p_2^{\,2}(\tau)\right]^{1/2}$

which, after a slight calculation, gives

$$W(t,p_1,p_2;\delta,\gamma) = t - \gamma + \delta + \int_0^{(\gamma - t)} \left[\epsilon^2(p_2 \sin\tau + p_1 \cos\tau)^2 + (p_2 \cos\tau - p_1 \sin\tau)^2\right]^{1/2} dt.$$

For a time optimal problem with autonomous system equations and target a a point in state space, the constant δ is inconsequential. We consider $\delta = 0$ and omit further reference to it.

By virtue of the transformation, solution trajectories to the system (3-15) with $v = v^*(t, x, p)$ are given by $x(t;\alpha,\beta,\gamma) = W_p(t,p(t;\alpha,\beta);\gamma)$ or specifically

$$x_1(t;\alpha,\beta,\gamma) = \int_0^{(\gamma - t)} \frac{\alpha\epsilon^2 \sin(2\tau+\beta)\cos\tau - \alpha\cos(2\tau+\beta)\sin\tau}{\left[\epsilon^2\alpha^2 \sin^2(2\tau+\beta) + \alpha^2\cos^2(2\tau+\beta)\right]^{1/2}}\, d\tau$$

$$x_2(t;\alpha,\beta,\gamma) = \int_0^{(\gamma - t)} \frac{\alpha\,\epsilon^2 \sin(2\tau+\beta)\sin\tau + \alpha\cos(2\tau+\beta)\cos\tau}{\left[\epsilon^2\alpha^2 \sin^2(2\tau+\beta) + \alpha^2\cos^2(2\tau+\beta)\right]^{1/2}}\, d\tau$$

These formulas can be interpreted as follows. If we choose $\gamma > 0$ and $t = 0$, $\left\{x(0;\alpha,\beta,0) : (\alpha_1\beta) \in E^2\right\}$ gives the set of initial points x_o from which the origin can be reached in time γ by trajectories which satisfy (3-15) with $v = v^*(t, x, p)$. In particular, it can be shown (via the theory of homogeneous contact transformations) that the jacobian determinant

$\frac{\partial (x_1, x_2)}{\partial (\alpha, \beta)}$ is zero, and in this case the set of initial points forms a closed curve in E^2 for each $\gamma > 0$.

To generate a field of extremals (it is to be cautioned that the term extremal is to be taken in the sense of the classical calculus of variations; i.e., not necessarily to infer optimality) choose $\gamma = 0$ and replace t with -t in (3-17). For each choice of α, β one obtains an extremal which is at the origin at time zero. Varying α, β now gives a field of extremals.

REFERENCES TO SECTION III

1. R. E. Kalman; The Theory of Optimal Control and the Calculus of Variations, Chapter 16 of Mathematical Optimization Techniques, edited by R. Bellman, University of California Press, Berkeley and Los Angeles, California (1963).

2. A. F. Fillipov; On Certain Questions in the Theory of Optimal Control, English Translation, J. SIAM on Control, Ser A, Vol 1, No 1, 1962.

3. A. E. Taylor; Introduction to Functional Analysis, John Wiley and Sons, Inc., New York (1958).

4. A. Auslander, R. E. MacKenzie; Introduction to Differentiable Manifolds, McGraw-Hill Book Co., Inc., New York (1963).

5. H. Buseman, Convex Surfaces, Interscience Tracts in Pure and Applied Mathematics # 6, Interscience Publishers, Inc., New York, (1958).

6. T. Bonnesen, W. Fenchel; "Theorie der Konvexen Korper," Chelsea Publishing Company (1948).

7. K. O. Friedricks; On the Differentiability of the Solutions of Linear Elliptic Differential Equations, Comm. on Pure and Applied Mathemathics, Vol VI (1953), pp. 299-325.